Submit a
LATE-BREAKING SESSION PROPOSAL
for JSM 2019

ALSO:
New Law Offers Reforms to Improve Access to Data, Confidentiality Protections

Editors of Online Journal *Statistics and Public Policy* Seek Papers

features

columns

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*.

This column is written for those interested in learning about the world of Data for Good, where statistical analysis is dedicated to good causes that benefit our lives, our communities, and our world. If you would like to know more or have ideas for articles, contact David Corliss at *davidjcorliss@peace-work.org*.

This column focuses on what statisticians do when they are not being statisticians. If you would like to share your pastime with readers, please email Megan Murphy, *Amstat News* managing editor, at *megan@amstat.org*.

Larry Lesser of The University of Texas at El Paso repurposed the KC & the Sunshine Band #1 hit "(Shake, Shake, Shake) Shake Your Booty" to concisely share major real-world examples in the history of random sampling: the 1936 Literary Digest poll and the 1970 draft lottery.

Random Sample

Lyrics © 2017 Lawrence M. Lesser

Two…. million readers

picked Landon to

advance;

Gallup

disagreed

from thousands picked by chance.

CHORUS:
(Shake, shake, shake,

shake, shake, shake:

random sample,

random sample) 2x

Ahhh, you think

you can pick them

very well,

But drafting

for Vietnam

wasn't swell.

(Repeat Chorus)
 Shake, shake, shake

shake, shake, shake:

iPod Shuffle,

quick pick lotto.

Shake, shake, shake,

take, take, take:

random sample!

MORE ONLINE
Inspired by the legacy and untimely death of Tom Short, Larry Lesser wrote a song in Short's memory, "Signal in the Noise." You can listen on YouTube at *https://bit.ly/2RTMYqW*.

CORRECTION
In the January issue, we inadvertently referred to Bradley Efron as being from UC Berkley. He is at Stanford. We regret the error.

departments

SDSS
SYMPOSIUM ON
DATA SCIENCE & STATISTICS
BEYOND BIG DATA: BUILDING DATA TOOLS
BELLEVUE, WASHINGTON • MAY 29–JUNE 1, 2019

Consider attending SDSS 2019 for a fresh look at your data research work – Page 25

member news

Follow us on Twitter
www.twitter.com/AmstatNews

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

Like us on Facebook
www.facebook.com/AmstatNews

Follow us on Instagram
www.instagram.com/AmstatNews

Statistics: Proactive or Reactive?

How to ensure statisticians decide the field's direction

The key word is *react*. The statistics profession's reaction to data science prompted my colleague to ask this question: "Why does our profession often find itself in reactive mode—reacting to others' direction—rather than being proactive and setting the direction ourselves?"

It's a good question. In some sense, it has been the basis of our profession for many years. During a special lecture at Indiana University in March 2008, Peter Hall said, "Statistics is 'reactive': It is very responsive to new problems that arise in chemistry, biology, physics, ..."

We've been fortunate to have advances in our field driven by others who bring us interesting data that prompt the development of new methodology. We did not have to create those new data types—others did, and then brought them to our attention. A great example of this took place during ASA President Jonas Ellenberg's invited address at JSM 1999 in Baltimore, when Eric Lander described the tsunami of genomic data desperate for statistical methods and research. (Others may have been introduced to these problems in other ways.) It was a call to action, to which many in our community responded, leading to meaningful advances both in statistical research and genomic science. It generated productive collaborations and energy in an inter-disciplinary field that we now call bioinformatics, which "combines biology, computer science, information engineering, mathematics, and statistics to analyze and interpret biological data" (Wikipedia!).

Bioinformatics presents a positive situation in which we stepped up to the plate presented to us. But do we have to wait for scientists to drop the problems on our laps? Or can we take another approach, one in which we create the opportunities—before others do and develop new disciplines and institutes to solve them before statisticians even realize what's happening?

On a tiny scale, statistics created its market in forensic science, but, again, it came about somewhat by serendipity. The National Academy of Sciences (NAS) Committee on Scientific Assessment of Bullet Lead Elemental Composition Comparison, formed in 2003, needed statisticians.

Most committees in NAS's Division on Earth and Life Studies never included a single statistician; this one couldn't avoid it because two of the four aspects of the charge clearly related to statistics. Cliff Spiegelman of Texas A&M and I were invited to participate, Cliff for his years of expertise with statistics in chemistry and chemometrics and I because of my knowledge of chemistry (from college, National Institute of Standards and Technology, and HP).

Seeing how much impact statistics had on that study, the study director for the Committee on Identifying Needs of Forensic Science Community, Anne-Marie Mazza, invited Constantine Gatsonis of Brown as co-chair and me as a member. The final report, *Strengthening Forensic Science in the United States: A Path Forward*, emphasized the importance of statistics in forensic science at a time when many forensic practitioners may not have known just what statistics is.

Maybe we can call that one pseudo-proactive: We moved on it, but only after statisticians were called in and, SURPRISE!, we added more value than our "customers" (the forensic community) expected.

Data science is another example, and statisticians are responding in different ways. Some are waiting to see if "data science" goes the way of expert systems and artificial intelligence—now seen as largely oversold—and continue to conduct statistical research needed for sound inferences. Some academic departments have changed their names to "Statistics and Data Science" in hopes of directing the data science students and consumers in our direction. And others are ready to abandon traditional statistics altogether and jump on the data science bandwagon.

David Donoho wrote in the *Journal of Computational and Graphical Statistics* article "50 Years of Data Science" that, "[T]he activities that preoccupied [the statistics profession] over centuries are now in the limelight, but those activities are claimed to be bright shiny new and carried out by (although not actually invented by) upstarts and strangers. Various professional statistics organizations are reacting ..." He cites three commentaries (Marie Davidian in *Amstat News*, July 2013; Bin Yu in *IMS Bulletin*, October 2014; and Marvin Goodson at the meeting of the Royal Statistical Society in May 2015). Three leaders in three leading societies in three subsequent years *react* to data science initiatives—long after these initiatives had taken hold in universities, government agencies, and corporations.

What are the messages to our field? Shall we avoid challenges and stay within our comfortable

Karen Kafadar

David Williamson's Impact Initiative committee is designed to address ways the ASA can inspire proactive activities for statisticians.

territory where our familiar approaches can be applied? (Recall the Tukey quotation at the end of my last column: "To be spared the responsibility of working on any of [these major problems] would make anyone's life simpler and more pleasant.") Or do we accept what Tukey called the "necessarily approximate nature of useful results" and that "indication procedures [may need] to grow up before the corresponding conclusion procedures"? In short, do we, as Tukey put it, "do what [we] can to clarify the issues and offer good advice … give up the protection of unquestioned hypotheses and contribute [our] best acuteness and wisdom under uncomfortable circumstances"?

Tukey advised us in *The Future of Data Analysis* that "data analysis is intrinsically an empirical science" and that, among other "necessary attitudes" for effective data analysis, we need to address "more realistic problems," accept the "necessarily approximate nature of useful results," and adopt the "free use of ad hoc and informal procedures in seeking indications."

Panelists at JSM 2018 in Vancouver (Session 149: Theory versus Practice) made some of the same comments: We can justifiably take pride in demonstrating the mathematical properties of proposed statistical approaches, but perhaps we also need a balance between the time to demonstrate mathematical validity and the potential delay in announcing our breakthroughs while we wait

for the "corresponding conclusion procedures" to "grow up." [Recall that the statistics community was introduced to the jackknife via an abstract, seven sentences long, in *The Annals of Mathematical Statistics*.]

Can we find morals in these stories?

a. We need to showcase all our talents—logical thinking, identification of process steps, design of relevant data collection, analysis and inference, characterization of uncertainty, clear results.

b. We need to not just seize opportunities, but create invitations to the table. Better yet, we need to create the demands for our talents.

c. Sometimes we need to be prepared to use our skills to present reasonable approaches to solve problems, even before the theory has been developed, before others offer less sensible approaches.

Dictionary.com defines proactive as "serving to prepare for, intervene in, or control an expected occurrence or situation" and reactive as "to act in response to an agent or influence; to respond to a stimulus in a particular manner." Most ASA members are too busy to be proactive. We're too busy *reacting* to our plates full of problems! That is why we have an ASA—to address issues of concern to the statistics profession that will ensure the relevance of our future.

Toward that end, David Williamson's Impact Initiative (*https://bit.ly/2HzSJWW*) is designed to address ways the ASA can inspire proactive activities. The committee he is chairing hopes to collect examples of fields or problems in which statistics had a critical impact (item (a) above) and, more importantly, identify whole areas where statisticians have the opportunity to proactively create the demand for statistical input, which will include analysis and inference of data, big and small (item (b) above). Please offer your suggestions at *https://bit.ly/2CEihwg*.

Martin Luther King Jr. was celebrated last month with reminders of a quotation from Mahatma Ghandi: "Be the change you wish to see in the world." What can the ASA do to proactively meet the next big scientific initiative? Thank you for helping us move our profession from "reactive" to *"proactive"* mode!

Sloan Foundation Funding to Enhance Data Science Expertise in Federal Government

With support from the Alfred P. Sloan Foundation, the ASA and its sister societies will sponsor three AAAS Science and Technology Policy Fellows in the federal government in each of the next two years.

The fellows sponsored by the ASA, American Mathematical Society, Association for Computing Machinery, Institute of Mathematical Statistics, Mathematical Association of America, and Society of Industrial and Applied Mathematics will have expertise in data science—particularly causal inference and machine learning—and serve one year in a federal agency or on the staff of a senator, representative, or congressional committee beginning September 2019.

"We are grateful to the Alfred P. Sloan Foundation for this opportunity to become sponsors of the prestigious AAAS Science and Technology Policy Fellowships," said ASA Executive Director and principal investigator on the proposal to the foundation Ron Wasserstein. "We wholeheartedly support their desire to bring more statistical and data science to the federal government to meet its administrative, legislative, and policymaking challenges."

"The ASA has a priority to raise the profile of statistics and data science in policy because of the many benefits that would result," said 2018 ASA President Lisa LaVange. "I'm thrilled for this opportunity. What better way to raise the profile than through people? I urge my colleagues in the statistical and data science community to apply. Having served in government and being proud of my accomplishments, I'm confident the statistics and data science fellows will make valuable contributions and find their service enriching and gratifying."

Since 1973, the AAAS Science and Technology Policy Fellowships program has connected science with policy, fostering a network of science and engineering leaders who understand government and policymaking and are prepared to develop and execute solutions to address societal challenges. AAAS places about 300 people each year through this program. ∎

Nominate a Colleague for ASA Fellow!

The designation of ASA Fellow has been a significant honor for nearly 100 years. Under ASA bylaws, the Committee on Fellows can elect up to one-third of one percent of the total association membership as fellows each year.

Individuals are nominated by their ASA-member peers. To be selected, nominees must have an established reputation and have made outstanding contributions to statistical science. The Committee on Fellows evaluates each candidate's contributions to the advancement of statistical science and places due weight on the following:

- Published works

- Position held with employer

- Membership and accomplishments in other societies

- ASA activities

- Professional activities

To be eligible for nomination, a person must be a current member of the ASA who has held continuous membership from March 1, 2016, to February 28, 2019.

Nominations may be submitted online until March 1 at *ww2.amstat.org/fellows/nominations/index.cfm*.

Statistical Leadership: Its Importance, Imperatives, and Impacts

A Summary of the 2018 Regulatory-Industry Panel Discussion

Steve Lane, Adeniyi Adewale, and Allison Florance

The 2018 ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop held in September featured a session titled "Statistical Leadership: Its Importance, Imperatives, and Impacts." The session—organized by Adeniyi Adewale and Allison Florance of Novartis and chaired by Steve Lane of BMS—consisted of two presentations representing industry and regulatory perspectives on statistical leadership by Gary Sullivan from Espirer Consulting and John Scott of the FDA, respectively. After the presentations, panelists Eric Gibson of Novartis, Gary Rosner of The Johns Hopkins University, and Janet Wittes of Statistics Collaborative corroborated the presentations with their own experiences. Here are summaries of the presentations and an excerpt of the discussion that followed.

Gary Sullivan

Sullivan defined leadership as the ability to deliver value consistently to an organization or cause by inspiring people to take a specific direction when they truly have the choice to do otherwise. He used examples from his experience in nonclinical statistics to underscore the following three key messages:

1. **The skills and insights that come easily to statisticians likely do not come as easily to colleagues in other functions, so don't hesitate to use those skills to add value.** In his early days as a nonclinical statistician and while engaging with a team on an issue dealing with variability of tablet potency, Sullivan collated data and did an analysis that provided important insight. He initially hesitated to share the analysis because it seemed "simple and obvious," but he soon realized his simple analysis held the key to resolving the issue.

2. **Statisticians have a perspective unlike anyone else on their teams and can use it to add value in a unique way.** While Sullivan was providing statistical support for a bio-product development group, he realized he was likely the only person in the department who worked closely with every group or function—analytical, molecular biology, microbiology, fermentation, engineering, purification, and pilot plant operations. This gave him a broad and deep understanding of the business, credibility with the entire organization, and a perspective no one else had, which allowed him to provide influence across the organization tactically and strategically.

3. **As statisticians advance to more senior levels and gain influence, they need to have more of an ownership mindset and contribute at higher levels.** Sullivan shared a situation in which he learned from a business partner that manufacturing deviations were creating a drain on time and resources. Many of these deviations were caused by out-of-control signals from control charts of product characteristics. Sullivan realized the control-charting, although well-intended, was creating cost and not providing insights or value to the business. He went through the process of shutting down most of the control-charting, thereby saving valuable time and resources that could be re-focused on troubleshooting and improvements.

Sullivan concluded his presentation with the following three points:

1. Statisticians have a strong foundation for leadership (problem solving, critical thinking, information processing, integrity), but must improve on decision-making, business judgment, and acting on ideas to be more effective.

2. Statisticians spend a significant amount of time improving their technical skills and relatively little time honing their soft skills, including leadership. Investing in the latter would serve them well.

3. More senior statisticians need to actively play a role in advancing the soft skills of less experienced statisticians.

John Scott

Scott opened his presentation by noting that any statistician can provide statistical leadership, the goal of which is to drive better decision-making. He underscored what leadership is about by borrowing the catchphrase, "Management is getting people to do what you want. Leadership is getting people to want what you want."

To drive better decision-making, statisticians must invest time into thinking about the problems at hand. Statisticians must refrain from reflexively doing what collaborators think they want and instead think about what the real scientific questions are and how they can be addressed with the resources and tools at hand.

He advised statisticians to give collaborators genuine help in generating, summarizing, and analyzing information and relayed an example of an engagement in which a laboratory scientist was interested in comparing antibody levels between three conditions and the differences were found not to be statistically significant. The scientist decided to collect more data and returned with the question, "Is it significant now?" Rather than address the question of statistical significance, Scott was able to structure the problem in a Bayesian framework, synthesizing the previous and new data. He noted that the write-up was fully transparent about the experiment and the sequential nature of data collection. The results were presented using descriptive statistics and credible intervals, not p-values or binary conclusions about significance.

The scientist did not get exactly what he asked for, but what he got addressed the real question. Statisticians must avoid being boxed in by methodology in approaching problems.

Scott disavowed the notion that "statistics is a fortress to protect us from marauding commercial interests" or "statistics is a way to say 'no,' never a way to say 'yes'." Physicians sometimes look for statistical justifications to establish preconceived notions.

To drive home his point, Scott shared an experience in which a regulatory clinical team objected to a sponsor's proposal regarding a noninferiority study design. The sponsor proposed to use a non-inferiority margin on an absolute difference scale while the regulatory clinical reviewers were insisting on a risk ratio scale. The FDA statistician was asked to help provide statistical justifications to support a risk ratio, which had lower power and, according to John, is generally less interpretable.

By going beneath the surface and probing to understand the clinicians' perspectives, it turned out the real issue for the FDA was the concern that sample size resulting from the sponsor's proposal was considered inadequate to characterize the safety profile of the compound. The clinicians and statisticians then agreed to accept the sponsor's proposal of an absolute difference scale, along with a stipulated minimum sample size for adequate safety evaluation.

From Scott's perspective, technical knowledge, disease-area knowledge, and good communication tools are imperatives for influencing wise decision-making.

Finally, Scott advised that statisticians must be assertive, speak up, and not be afraid their opinion may be wrong or challenged. They must value their own expertise and literally sit at the table. Sitting at the back of the room conveys the message the statistician has no important contribution to the discussion.

Panel Discussion

Janet – Both Gary and John's talks addressed what leadership ought to be. Beware, however, of negative leadership. I once worked for two aggressive competing managers who both wanted me to fail so they could use my failure as a weapon against each other. In such counter-productive set-up, removing yourself from the environment (for example, quitting) is always an option. As you establish yourself as a leader, your role should be to encourage and help others succeed—that will increase the chance of success of the project on which you are all working.

> Be mindful of your biases and starting point. Do not approach a discussion with a mindset that your clinical partners do not want to listen to you; they are actually thirsty for your ideas and input.

Eric – Leadership is about influence without authority based on the relationship capital you have built over time working with people. One of our biggest responsibilities is cultivating the next generation of leaders by opening doors and creating opportunities for others. Don't make the mistake of evaluating future talent through the lens of your own experience; they are readier for opportunities than you may think.

Gary R. – There is a difference between statistical leadership and leadership in the broader sense. Everyone should take advantage of ASA leadership training, as not all places of employment offer formal leadership training. Relationship with clinicians can become confrontational because they come with preconceived opinion. It is our job to question and come to a mutually agreeable solution.

Watch the YouTube cartoon about consultation with a basic scientist (*https://bit.ly/2MbBE3S*).

How have you been able to shift someone's opinion?

John – Be aware of the personalities you are dealing with, speak with authority, and don't be intimidated.

Steve – It can be a simple act of engagement and conversation.

Janet – Statisticians really need to know where the individual expressing an opinion is coming from. Do not be afraid to say, "I don't understand." The scientist to whom you are talking wants to teach; take advantage of that!

Gary S. – Work on building the relationships and establishing trust. Meet with individuals to understand what they are doing and what challenges they are facing. Most physicians and scientists like to talk about what they are doing, so

you shouldn't have much trouble learning about the problems they are working on and what their interests are. Start by contributing, helping, and adding value to what they want before trying to redirect them or their work. This will demonstrate competence and build credibility. With patience and time, you will be able to interject, have your opinions be heard, and have influence with them and their work.

Eric – Be mindful of your biases and starting point. Do not approach a discussion with a mindset that your clinical partners do not want to listen to you; they are actually thirsty for your ideas and input. Step out of your comfort zone and talk to people. Statistics is also about helping scientists find the right question, and that takes legwork. Build relationships, ensure two-way conversation. Influence is process and not a single event, and it is depends on the network of relationships you develop with those you are seeking to influence.

Gary R. – Working together among ourselves and with junior-level statisticians to develop courage and confidence. Lead by example.

Audience: How can we breach the difference between academic training that emphasizes independent contribution and the industry reality of collaboration and interdependence?

Gary R. – Academic statisticians must continue to advocate for elements that are missing across academic programs. Some schools have started offering courses useful for statistical leadership. Students should be encouraged to seek internship opportunities.

Eric – Going back to my experience in graduate school, it was helpful to be able to shadow someone in academic consulting engagement. Also, in industry, bringing junior team members to the table is important for their development so they can see the dynamics of the decision-making process.

Janet – As Eric stated, senior statisticians should provide opportunities for junior statisticians to shadow them. Statisticians should not be afraid of saying something stupid. Watch others; they are not afraid of dumb comments. We often keep silent because we fear looking ignorant.

Gary S. – Teaching leadership courses in graduate school is one thing that will help. There has been a course taught at UNC since 2010 that was co-developed by Lisa LaVange. I am talking with Purdue and may be teaching a leadership course there in 2019.

John – Formal training in soft skills is also important for leadership development. Coursework

or a practicum in teaching or even in making PowerPoint slides could be more important in the long run than a lot of the technical coursework currently in the curriculum.

Audience: Physicians desire to make decisions, while statisticians think they should only present the facts. How do we define the scope of statistical leadership?

Gary S. – It's important to begin to set expectations for statisticians early in their careers to see themselves as leaders. Just as they will continue to invest in developing their technical skills, they also need to invest in their leadership skills.

Eric – It is exciting to work in an organization in which the CEO places value on statistics and quantitative sciences. In an organization like that, statisticians should continue to think about where they can add value to the enterprise. Constantly ponder the question, "How can we continue to engage and be broadly relevant to the problems confronting the organization?" We have an example in Novartis in which statisticians have collaborated with other functions in the organization to develop tools to facilitate quantitative decision-making on drug development portfolios. Statisticians should speak up about their ideas and not be afraid of some ideas not succeeding.

Audience: How do we fight the perception that statistics is a narrow discipline?

John – Statisticians should take the opportunity to engage in the subject matter in which they are working. It is a privilege to be in a narrow field that can affect everyone else's discipline. I encourage people to read about the medical area they are working in; even starting with a Wikipedia article can be a big help.

Janet – It is not bad to be an expert in a narrow area. One can be an expert in a narrow area but have broad interests and participate in a broader multidisciplinary space.

Gary S. – Statisticians should leverage the insights of supervisors and superiors on development opportunities. If you have an interest in advancing administratively or technically, you need to have the conversation with your supervisor or your functional lead. Ask questions such as "What skills do I need to develop to be considered for that position or promotion?" "Can you help me find an opportunity to gain pertinent experience?" It is the responsibility of a supervisor to have those discussions with you and provide you appropriate feedback and direction. ■

Independence of Puerto Rico Institute of Statistics at Risk Again

Puerto Rico's governor has appointed three new members to the Puerto Rico Institute of Statistics.

Puerto Rico Governor Ricardo Rosselló's appointment of three new members (see *https://bit.ly/2VURDI7* [Spanish]) to the board of directors for the Puerto Rico Institute of Statistics (PRIS) in November has again raised concerns about the agency's independence. According to advocates for transparency, open government, and evidence-based policymaking, two of the appointments violate the law that established PRIS.

The law states members of the board must be experts in fields related to statistics, mathematics, economics, or planning. Also, only one member of the board can be a government official. One of the newly appointed members lacks the specified credentials, while another is a government official, which exceeds the limit of one.

Earlier in the year, Governor Rosselló proposed dismantling PRIS by consolidating it into another government department, removing its independence protections, and outsourcing its statistical research. That plan followed the governor's illegal removal of four PRIS board members in 2017. PRIS was preserved after a last-minute intervention by the Puerto Rico Senate.

With the July news of PRIS being preserved, 2018 ASA President Lisa LaVange stated, "Having independent, strong, and transparent statistics will serve the health, well-being, and economy of Puerto Rico and its people for decades to come."

The American Statistical Association is concerned about what appears to be a renewal of the attempts to undermine the Puerto Rico Institute of Statistics and its independence. The ASA urges the laws protecting the independence of PRIS be followed and will work to ensure they are. ■

Mathematics and Statistics Awareness Month: What's on the Agenda?

Lara Harmon, ASA Marketing and Online Community Coordinator

There's more to look forward to in spring than seeing the sun again! Every April, math and statistics associations across the country celebrate Mathematics and Statistics Awareness Month, a full month of programs and initiatives dedicated to promoting better public understanding of mathematics and statistics.

The ASA has big plans for April this year. We're hosting our first virtual science fair, a chance for middle- and high-school students to put their problem-solving skills to work addressing real-life issues. Students will tackle the theme "Solving Problems with Statistics" as they choose, collect data on, and work to find possible solutions for a problem facing their community.

While students prepare for the virtual science fair, educators can look forward to four live Q&A sessions with professionals who make a living solving problems using statistics. Our guests will introduce students to problems surrounding the following:

- Water quality and public health (Loni Tabb, associate professor of epidemiology and biostatistics at Drexel University)

MSAM
Mathematics & Statistics Awareness Month

- Machine learning and video game AI (Eric Laber, associate professor of statistics at North Carolina State University)

- Understanding athletes' performances (Stephanie Kovalchik, research fellow at the Institute of Sport, Exercise, and Active Living, Victoria University)

- Caring for the environment (Erin Schliep, assistant professor in the department of statistics at the University of Missouri

While watching the prerecorded sessions, students will be able to ask questions in real time via chat and receive live answers from our guests. We'll announce the schedule for these later in the year, but expect one for each week in April.

Of course, you can also look forward to our annual Mathematics and Statistics Awareness Month poster in the April issue of *Amstat News*. ∎

USA Science & Engineering Festival Team Expands Efforts to Cities Outside DC

The success of the annual X-STEM symposium in Washington, DC—an event featuring interactive presentations from STEM visionaries—has led the USA Science & Engineering Festival team to branch out to other cities in the United States beginning with X-STEM at Travis Air Force Base in Northern California this month.

X-STEM DC, presented by NCR Foundation, will return March 19, with more than 3,000 students ready to engage with real-life STEM role models. Learn more at *https://bit.ly/2VSVCoH*.

Next up is the Clippers SciFest SoCal March 22–23 at the Los Angeles Convention Center. The goal of this event is to inspire a new generation of STEM professionals on the West Coast, tapping into the LA Clippers' strong relationships with businesses, schools, community groups, and government stakeholders. Clippers SciFest SoCal will start off smaller in size than the DC Expo, but will incorporate all the components. Check out *https://bit.ly/2Fv3s2a* for details.

With the theme "Vision for STEM," the 2020 expo is set to return April 25–26, 2020, to the Walter E. Washington Convention Center as the nation's largest STEM education and workforce development event. More information is available at *https://bit.ly/2VRof5D*.

For more information about the USA Science & Engineering Festival, visit *https://usasciencefestival.org*.

Editors of Online Journal *Statistics and Public Policy* Seek Papers

Jerry Reiter and Mike Cohen

Statisticians and other quantitative analysts often carry out evidence-based research that has direct relevance to important public policy questions. Such research might involve, for example, careful modeling of data to address issues of causal inference, investigating the properties of methods typically used in policy evaluation, or explaining the results of elections.

The ASA journal *Statistics and Public Policy* publishes peer-reviewed articles describing such research on outstanding, complex public policy issues, where the use of appropriate statistical methods is shown to clarify such questions (*https://bit.ly/2sucyEl*).

Statistics and Public Policy (*SPP*) is an online, open access journal. The editors are looking for articles that apply sound statistical thinking and methods to issues related to public policy. Unlike other ASA journals, submitted articles need not include innovative statistical methodology—insightful and appropriate application of existing techniques is perfectly acceptable. Indeed, this is the sweet spot for *SPP*. For example, we welcome submissions in which the statistical methods may not be quite innovative enough for publication in the *Journal of the American Statistical Association* or the *Annals of Applied Statistics*, but are perhaps more involved than what is typically published in substantive journals. We encourage submissions from all areas of statistics and public policy, including education, energy, the environment, health, politics, and policy evaluation.

The audience for the journal is interdisciplinary and large. Some articles have readership counts approaching 10,000.

Recent articles include the following:

- "Inference of Long-Term Screening Outcomes for Individuals with Screening Histories," by Dongfeng Wu, Karen Kafadar, and Shesh N. Rai uses a patient's screening history to assess future likelihood of disease and the benefits of future screening.

- "A Spatial Study of the Location of Superfund Sites and Associated Cancer Risk," by Raid Amin, Arlene Nelson, and Shannon McDougall, assesses the impact of the pollution associated with superfund sites on cancer incidence.

- "Study of Salary Differentials by Gender and Discipline," by Lynne Billard, examines how the use of models that failed to include interaction effects could mislead about the impact of gender and discipline on academic salaries.

- "19 Things We Learned from the 2016 Election," by Andrew Gelman and Julia Azari, offers and evaluates a variety of explanations for the results of the 2016 presidential election.

- "Officer Risk Factors Associated with Police Shootings: A Matched Case-Control Study," by Greg Ridgeway, features how the officers were linked to the use of deadly force.

SPP editors seek to increase the number of published articles and add special features. For example, they are planning special sections dedicated to statistical methods for characterizing partisan gerrymandering when drawing electoral districts, articles describing the US federal government's data strategy, and possibly articles looking into the details of the most recent midterm elections.

You can submit manuscripts for peer review using the Scholar One management system. Links are available on the *SPP* main web page at *https://bit.ly/2sucyEl*. If you have questions about the journal, contact the editor, Jerry Reiter, at *jreiter@duke.edu*. ∎

Developing a Privacy Policy

ASA Privacy and Confidentiality Committee Members

There is some risk of disclosing identities of individuals (or entities) and their sensitive data that exists in every data file released. A privacy policy should describe a process that considers whether the release of the data file followed reasonable procedures and a process that applies appropriate data use restrictions and access controls. At the very least, the risk of re-identification should be minimized after considering the possible approaches to re-identify persons using existing technology and public or commercially available information.

Furthermore, it is important to be aware that the privacy of individuals can be affected by a data release whether or not the person is in the file.

The European Union (EU) recently adopted General Data Protection Regulation (GDPR), which has raised awareness of maintaining data confidentiality. With the adoption of GDPR and more awareness of data confidentiality, organizations should include the GDPR in their privacy policy as applicable. For more information about the GDPR, see the July 2018 issue of *Amstat News* at *https://bit.ly/2O8UDjU*.

Background

In the US, the legal framework for privacy policies is based on various federal and state statutes and case law. The majority of these laws focus on protecting the identifiability of persons as the basis and approach for protecting privacy interests. Privacy interests are fragmented across subject matter topics, and privacy protections and remedies also vary across subject matter.

The current requirements to protect privacy interests place a strong emphasis on whether personally identifiable information (PII) exists in the data file. The definition of PII is so broad that it refers to any piece of information that can be used on its own or with other information to re-identify, contact, or locate a person in a file.

The reliance on a vague term such as PII for whether a data system requires privacy protection provides no guidance for developing a privacy policy in the digital age. The reliance on the presence or absence of PII in a data file should not be determinative on whether privacy protections apply.

Recommendation

The current privacy legal framework needs to change and focus on applying appropriate data use restrictions, assessing the risk of re-identification, creating the process for developing the data set, and determining the contents of a data set.

A privacy policy should focus on ensuring the process and procedural actions to protect information are followed. New statutes need to be drafted that require privacy protections regardless of whether PII exists in a file and impose penalties without requiring persons to show harm or damages.

The new legal framework needs to focus on data use restrictions and access controls as the main approach for protecting privacy interests throughout the lifecycle of information, beginning with data collection and ending with record destruction.

Guidance

Privacy interests should apply, regardless of whether persons are identifiable in a file and regardless of whether PII exists. The existence of PII in a data file is only one factor as a high two-risk variable in determining the identifiability of the data. For example, combinations of indirect identifying variables can lead to the identity of an individual, or the full or partial data vector for a record may be linkable to another file containing PII.

A privacy policy should associate the levels of access and the level of data protection methods (data protection methodology) to minimize re-identification risk. For example, a data set in a physical enclave with strong access controls may be able to allow access to more raw data than a publicly accessible data set.

Strong privacy protections rely on applying appropriate data protection methods, data use restrictions, and access controls. Privacy rights should be enforceable without the need to show direct or indirect damages.

Privacy violations should be noted when the process described in a privacy policy is violated, the required actions are not applied, or the actions taken do not comply with the process.

Essential elements for a data privacy policy should, at the very least, do the following:

1. Describe allowable uses of the data. Some uses are more problematic and require more controls or restrictions. Determine whether researcher access to microdata will be allowed. Use specific informed consent statements and avoid the use of broad consent forms.

2. Describe the source(s) of the data and, to measure re-identification risk, review the sample size and whether sample weights are used in the design, the population size, whether high-risk variables exist in the data file that can be matched to external files, and whether multiple records in a data file are known to belong to the same cluster. Longitudinal and panel surveys create a special case of disclosure risk that

Journal of Privacy and Confidentiality Relaunched with Special Issue in Honor of Stephen E. Fienberg

Aleksandra Slavković and Lars Vilhuber

In addition to Stephen Fienberg's numerous contributions to the research, education, and practice of statistics, social sciences, and machine learning, he was a co-founder of the *Journal of Privacy and Confidentiality* (*JPC*) and—to the end—editor-in-chief of it. His absence led to a significant hiatus in the journal's activities.

Stephen Fienberg

The newest issue relaunches the journal. Aleksandra Slavković and Lars Vilhuber introduce the issue, including its special features, in their editorial. Additional editorials by Cynthia Dwork and Vilhuber address what the editorial team has accomplished in the past year and aims to accomplish in the next few years.

The issue is in honor of Fienberg and his impact on the methodology and practice of privacy and confidentiality. It commemorates the intersection between statistics, computer science, privacy, and confidentiality as he envisioned so many years ago. Six peer-reviewed articles, as well as 10 reminiscences about Fienberg by researchers who worked with him, can be accessed at *https://bit.ly/2srdOZ5*.

JPC is an open-access journal, and all articles are free to view and download.

JPC accepts submissions from any field as long as it relates to privacy and confidentiality. Details can be found at *https://bit.ly/2FqRLdv*.

---

may be associated with linked files. In this case, the disclosure risk of a microdata file increases if some records on the file are released on another file with more detailed or overlapping recodes (categorizations) of the same variables.

3. Describe the key re-identification risk elements and potential attack scenarios the policy is designed to protect against.

4. Determine what data protection methods should be applied to data releases that are commensurate with the proposed use(s).

5. Describe the data access controls for the data system. Reference whether any privacy protection certificates/certificates of confidentiality apply that reduce the risk of required disclosures of identifiable information.

6. Verify the organization has the resources, expertise, and capability to provide appropriate guarantees, assurances, or attestations for privacy protection when information is being collected.

7. Avoid broad consent statements and develop specific consent and use statements for the information collected.

8. State whether information will be disclosed to the third parties for other purposes such as research and/or marketing.

9. Describe the baseline anonymization process and determine the threshold level for re-identification of data subjects in the file to be publicly released. Consider how to minimize what information is collected and stored to reduce the probability of linking records to re-identify individuals.

New privacy statutes should provide clear enforcement sanctions via civil damages and criminal penalties for violating the requirements and procedures stated in the policy without the requirement for a person to show harm or damages. The important steps to follow for implementing a privacy policy include assessing the data to be released and the risks of disclosing identifiable information, minimizing the data to be released without compromising utility, applying reasonable data protection methods, developing and applying appropriate monitoring, and creating an accountability and breach response plan (i.e., use checklist or process approach similar to the approach used for developing a privacy impact assessment).

Data Release Policies

A data release policy should be consistent with and support the privacy policy. Factors to consider when drafting a data release policy include the following:

1. Different approaches for assessing re-identification risk

2. Different assumptions for determining the sophistication of the data intruder

   a. The likelihood the intruder would attempt to re-identify. The likelihood is often unknown, so assumptions need to be stated regarding the sophistication and competency of statistical and computational skills needed for re-identification. Apply a reasonableness standard in terms of technology, software, and

auxiliary data files that the data supplier is aware of after making a good faith effort to identify other linkable data files. Auxiliary files that may not appear useful for linking records may have greater value in the future, depending on the growth of available data sources. This requires an annual review to evaluate risk and account for an increase in available data files within a global network and the development of new statistical skills and methods.

    b. The amount of effort an intruder would spend trying to re-identify. There may be facts or circumstances the data supplier is aware of that may change these assumptions and require greater protection, such as the likelihood of linkable auxiliary data that can be used for re-identification or the intruder being a foreign government or other entity with significant resources available.

3. Different types of harm that can arise from an unauthorized disclosure

4. Differences in the resource capability, expertise, and effort an organization can spend for applying and testing data protection methods

5. Differences in the utility of the de-identified data; in evaluating the utility of the data, the needs of researcher access to the actual data versus using synthetic data should be considered

The process for releasing data should include the following:

1. An enforceable commitment not to re-identify

2. An audit trail on access to data

3. A measure of re-identification risk before and after data protections are applied

4. Application of data protection methods to block reasonable efforts to re-identify persons

5. An anonymization process that applies to all data and follows a 'reasonableness' standard

6. Other access and disclosure controls such as federal privacy certificates, data licensing terms, or memorandum of understandings for enforceable control of data releases

No single risk policy is appropriate for all types of data releases. The minimum risk tolerance should be established on the probability of re-identification of a person in a data file. More broadly, the risk ought to be the disclosure risk of an individual or the sensitive characteristics, regardless of whether a person is in the database. ∎

February Issue Offers Variety of Topics

The February 2019 issue of *The American Statistician* (*TAS*) features 13 articles.

The General section begins with an article that delves into why p-values are controversial. The discussion centers on a connection between p-value and formal decision rules. The second article discusses marginal least squares estimators, which are important for the analysis of high-dimensional data. Asymptotic properties of these estimators are investigated under a correlated errors assumption. A third article revisits how to construct a confidence interval for a normal mean with just one observation when the variance is unknown. The article is dedicated to the memory of Charles Stein. The final article in this section discusses statistical characterizations through the use of equal-in-distribution identities.

That Statistical Practice section includes three articles. The first discusses desirable properties for Bayes factors, and then provides a detailed examination of different Bayes factors that have been proposed for a two-sample comparison context. The second article discusses the impact and wisdom of adjusting comparative analyses for observed covariate imbalance. The final article in this section compares the Wilcoxon-Mann-Whitney test with modifications of that test motivated from interest in H0: P(X>Y)=1/2.

There are three articles in the Teacher's Corner. The first details an interactive classroom exercise that can help teach the meaning of statistical confidence intervals. A second article provides naturally arising contexts where random sums of random variables arise and extends formulas for the mean and variance of these sums by relaxing independence assumptions. The third article is an experience report of a two-semester course set up to specifically teach statistical collaboration skills.

The issue has single articles in the Data Science and Statistical Computing and Graphics sections. The data science article is about clustering and shows how silhouette indices can be extended to provide probability-like measures of cluster membership. The computing article discusses replication challenges when software package implementations of multiple imputation procedures are used.

The issue concludes with an article in the History Corner that traces the origin of stem-and-leaf displays.

Visit *www.tandfonline.com/loiutas20* to learn about *TAS* and how to submit your work for publication. ∎

New Law Offers Reforms to Improve Access to Data, Confidentiality Protections

Nick Hart

I n December 2018, Congress paved the way for improved accessibility and protection for the data government collects. Through the Foundations for Evidence-Based Policymaking Act, Congress endorsed a suite of major bipartisan data reforms that will likely shape statistical policy for decades. The president enacted the provisions into law in January 2019.

The law takes steps to implement half of the unanimous recommendations of the 15-member US Commission on Evidence-Based Policymaking, which issued its final report in 2017.

The commission—which included former heads of statistical agencies, researchers, and program administrators—concluded that the federal government could take meaningful steps to strengthen privacy protections while also improving secure access to confidential data for statistical activities. Thirty-six former heads of federal statistical agencies endorsed the commission's recommendations, joining more than 100 organizations that issued similar endorsements.

Enactment of the legislation is a promising indicator about interest in reinforcing the key federal legal protections that exist for confidential data. By reauthorizing the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) as part of the legislative package, Congress emphasized the importance of the critical CIPSEA protections for the American public and facilitating public trust in government's statistical activities.

For the statistical community, the CIPSEA reauthorization also included new opportunities to facilitate secure access to data, expanding existing CIPSEA authorities. The law directs agencies to presume administrative data collected by government are available for statistical activities, unless expressly prohibited by law, and creates a common portal for researchers to apply for access to restricted data sets. These expanded authorities are coupled with increased responsibilities for statistical agencies to assess the risk individuals can be reidentified when statistical or information products are publicly released.

The law also includes provisions to improve coordination of key data management activities in federal agencies. When implementing the legislation in coming months and years, agencies will be encouraged to recognize the expertise of

In January 2019, a new law offered many bipartisan data reforms that will likely shape statistical policy for decades.

statistical agencies and establish new chief data and evaluation officers to provide senior leadership for responsible data use. Agencies will also be expected to make more data publicly available when possible, providing accessible data inventories and improved metadata.

While the Foundations for Evidence-Based Policymaking Act does not establish the commission's headline recommendation for a National Secure Data Service to facilitate activities related to temporarily combining data from across government agencies, the law does create an advisory committee to support and plan for efforts to establish such a data service. The committee is designed to be representative of a cross-section of data users for identifying potential implementation steps and privacy safeguards as discussion of the data service progresses.

In enacting this important law, Congress and the president set the stage for major reforms across government agencies in using and protecting sensitive data. Doing so will have tremendous benefits for the American public and policymakers' ability to understand and evaluate government programs and policies. Coupled with ongoing efforts to develop a long-term federal data strategy, the federal government is moving toward recognizing and using data as an increasingly valuable asset. ■

Nick Hart is the director of the Evidence Project at the Bipartisan Policy Center, where he leads efforts to support implementation of the recommendations from the US Commission on Evidence-Based Policymaking.

Using Big Data in Precision Medicine

Jonathan Chainey, Roche; Victoria Gamerman, Boehringer Ingelheim Pharmaceuticals, Inc.; John Quackenbush, Dana-Farber Cancer Institute; Cecilia Schott, AstraZeneca; Jane Wilkinson, Broad Institute; Marc S. Williams, Geisinger; Kelly H. Zou, Pfizer Inc.

The ultimate goal of precision medicine (PM) is "right patient, right medicine, right time." To identify the "right patient," both cross-sectional and longitudinal real-world data (RWD) are useful. Examples of RWD include electronic health records (EHRs), claims, patient reported outcomes (surveys, preferences), genomic, images, and laboratory.

Recently, the US Food and Drug Administration (FDA) established a strategic framework (see *https://bit.ly/2AVST5d*) to advance the use of real-world evidence (RWE) to support development of drugs and biologics.

Here, several experts from pharma, biotech, government, academia, and technology providers discuss the opportunity to use such big data to find the right drug at the right time using the right target, which is ripe for forming collaborative partnerships.

To encourage collaboration among diverse organizations, what is the high-level strategy as an overview? What can our resources achieve, and how do we need to partner in terms of value and culture?

The key to collaboration is establishing a shared vision and compatible cultures. RWD, typically in the form of big data, are characterized by big data's five Vs: volume, velocity, variety, veracity, and value (see *https://ibm.co/18nYiuo*). A value proposition has to be developed that provides value for all collaborating partners, which is the fifth V of big data.

To do so, the need for the collaboration should be identified across all involved parties, including the elimination of the belief that the same deliverable can be achieved by any single party involved. Value should be placed on the synergy of the collaborating partners working together such that the sum of the knowledge and insights gained from the collaboration is greater than if each were to pursue a piece on their own.

Value should also be placed on, to the degree possible, standardizing both data and data capture, as we are often faced not with data that are overly "big" (relative to data in other domains), but rather incomplete and "messy."

Even with well-standardized data, the challenge is synthesizing these diverse data sources in the context of a given clinical scenario. To develop the "right medicine," especially in disease areas such as rare diseases, administrative claims data may help assess and optimize health care providers'

therapeutic decisions, monitor adherence, assess gaps in therapies, and evaluate switches among different dose levels and therapeutic options.

The health care industry puts patients first to drive innovations and develop the most appropriate medicines and treatment courses. To capture the "right timing" for treatment, both real-time instantaneous and longitudinal data are necessary.

Based on collaborative research to advance precision medicine with -omics and phenotype data, can you describe some specific case examples where there were successful partnerships?

For example, the collaboration between Geisinger and Regeneron has been positive. Various organizations were able to work together to develop a shared vision of combining genomic data with electronic health record data to support discovery, clinical care, and population health.

The value proposition demonstrated success for Regeneron by allowing them to develop new drug targets at a lower cost and for Geisinger by making available large amounts of genomic data that can be returned to patients to improve care. Both partners have had success in discovery research leading to high-impact joint publications.

With the shared vision in mind, long-term sustainable collaborations also result from a clear understanding of each partner's currency (i.e., what is important to the organization) throughout the lifetime of the collaboration, which can best be achieved from pooling complementary cross-functional expertise and resources.

Considering operationalization and implementation, how can we implement incentives that encourage clinical data to be standardized and shared?

Progress is being made in the standardization of clinical trial data within the pharmaceutical industry, where the regulators play a key role. Specifically, since December 2016, the FDA has mandated trial sponsors submit data in accordance with data standards developed by the Clinical Data Interchange Standards Consortium (CDISC) as part of the electronic submission process for regulatory approval. Data in CDISC format is now also accepted by the PMDA in Japan, which will be mandated in 2020. Other major regulators across the globe also endorse CDISC.

While there is still progress to be made in consistency of implementation, CDISC is uniquely

positioned to drive this international standardization of clinical trial data, and the progress it has made to date would simply not have been possible without the 'regulatory incentive' for trial sponsors to adopt it.

While closely related, data sharing presents a unique set of challenges, with patient privacy and patient consent of paramount importance. To encourage and require the incentives in place for trial sponsors to share their clinical trial data, regulators must play a crucial role.

For example, the European Medicines Agency Policy 70 will lead to the publication of redacted clinical study reports. Such policies can lead to win-win scenarios in which trial sponsors can access each other's data or researchers have access to trial data—provided those researchers agree to share their own insights with the trial sponsor.

This more open approach has started to bear some fruit, as can be seen through projects such as consortiums like ClinicalStudyDataRequest.com (CSDR) and Project Data Sphere.

In collaborations, the incentives to standardize data for sharing are evident. Negotiations around which standards to use for the purposes of the collaboration can yield an agreed-upon strategy. The challenge is creating incentives to drive a national or international standardization around these data.

As more use cases emerge, and with the work of groups like eMERGE and GA4GH, we anticipate more useful and universal standards will emerge in the near future. That will provide the foundation needed to really accelerate the use of these data in research and clinical care.

Big data and machine learning/artificial intelligence are, in many ways, no different from quantitative and statistical methods that have long been used in health and biomedical research and operations research. All these methods attempt to fit available data to predictive models that ultimately can help improve performance and outcomes of the system.

What has limited past and present approaches to these problems hasn't been the lack of methods, but the lack of outcomes data in sufficient quantities to build useful models. When we consider the five Vs of big data, the increases in volume, velocity, and variety in health care are obvious. The release of more comprehensive outcomes data is what will lead to value, since outcomes are what we need to match the right patient to the right drug. And it is continued access to outcomes data that will help us assure the veracity of our data and the models we build on them.

We clearly see the potential value of big data in building precision medicine. It is up to all of us

Getty images

to work together to see this potential realized to advance science, drug discovery, clinical practice of medicine, and—most importantly—outcomes for the patients we care for.

How are the regulatory landscape and directions in the US or globally incorporating big data into their approval process?

RWE was defined in the US 21st Cures Act (RWE, RWD, and big data; see *https://bit.ly/2D9iPvj*). RWE has received considerable interest, with great potential in health care policy and data science. Large volume is only one aspect of the data providers and patients deal with. Thus, the right analytic strategies will require increased resources and expertise.

The path from big data to precision medicine was recently discussed by a group of panelists when sharing best practices/experiences on fostering a collaborative approach to evaluating -omics and phenotype data. Geisinger's project is embedded in a real-world health care delivery system. The project includes a collection of clinical data from the EHR, supplemented by patient-reported outcomes to collect evidence and assess the value of return of genomic results.

Moving away from the traditional blockbuster model, how does PM present new opportunities for collaborations in the health care industry?

PM relies on a patient's genomic data and RWE. Big data analytics and artificial intelligence technologies are the key to unlocking the power of clinical data and thereby accelerating clinical development. Realizing the potential of and effectively using this data can mean the difference between a response and a failed trial. Big data has potential to drive new insights, but the ultimate value will need to be discussed by the stakeholders involved in the collaboration to advance precision medicine. ∎

So, You've Decided to Study Data Science … Now What?

ongratulations, you just made a choice that will bring you a fancy-sounding career! Predicting elections, forecasting business growth, identifying people in risk of cancer… a future of working on these types of projects is motivating.

As a Duke senior who has had a long journey studying data science and is going to a tech company in San Francisco after graduation, I want to share with you—aspiring freshmen, sophomore, and junior data science students—some insights I wish I'd had earlier.

Imagine we are grabbing coffee and you ask me, "What should I do to be a data science (DS) wiz after undergraduate school?" Here are the three points I would make to you:

A data science candidate needs to have a combination of skills, but you don't need to be the master of everything.

It ranges from data wrangling and statistics to business sense to interpersonal skills. It's hard because

we need to be more well-rounded and proactive in learning certain skills not taught in our undergraduate courses. It's easy because it's so fun and so powerful when we learn these skills.

Drew Conway invented a DS Venn diagram that indicates the section where DS lies (Figure 1).

From my experience, we can view a new graduate DS job as that of a consultant who solves problems by coding up data analytics. We don't need to be as strong of a coder as a software engineer or as talkative as a management consultant. But some coding skills plus business/product sense, in addition to a great understanding of statistics, will make us excel in DS.

You don't need to have a DS title to do data science.

"Data scientist" (DS) sounds fancy, while "data analyst" (DA) sounds cheap, so a DS is definitely better than a DA, right? No.

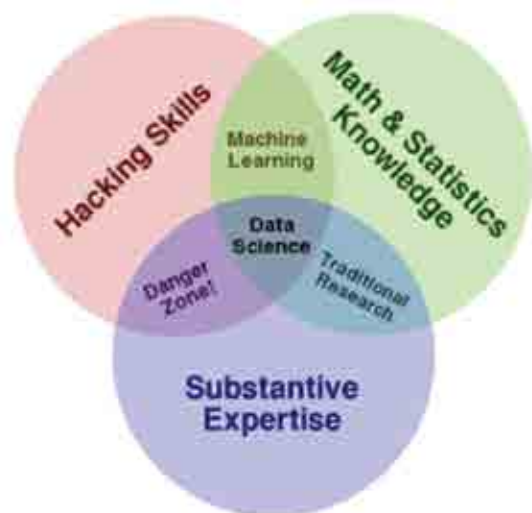The whole DS industry is so young, and the term "data science" is not well defined. In one

Figure 1: Venn diagram by Drew Conway showing where data science resides

company, a DS could be a PhD with five years of working experience building production-level model pipeline for fraud detection. In another company, a DS could be a new graduate doing charts in Excel.

By the same logic, when you job search, do not only look for jobs with the title "data scientist." Sometimes, the fancy jobs that harness the skills of machine learning, data wrangling, and business recommendation have titles such as "data analyst," "product analyst," or "business analyst," while data scientists are those with advanced degrees and industry experience.

In my case, I turned down a DS position from Company A and signed on as a product analyst with Company B because a product analyst at Company B was the more intellectually challenging position.

So, don't just look at the title. Make sure you read the job description and talk to people inside to learn what it is like to work at a particular place.

What should I do to prepare for becoming a DS after college?

1. You should take more applied classes in statistics and computer science (electives and graduate-level courses). As a stats major at Duke, I found that while the undergraduate statistics core classes laid out a strong foundation for statistical theories, it was the electives and graduate-level courses that actually helped me in interviews and carrying out projects in a company. I recommend taking classes that teach you how to code in R / Python for statistical computing and classes that teach you how to build models and do machine learning. In interviews, you will be tested on your data wrangling skills and statistical modeling knowledge, but rarely on deriving gamma distributions.

2. You should develop your business/product sense. Doing DS (and not software) means you will not be coding 100% of the time. It is a job that also involves strategy, problem solving, and presentation. Since sophomore summer, I've been taking business classes, subscribing to business/tech email newsletters, and listening to business podcasts. The days before I flew to interviews, I spent hours reading up on those employers; testing their products; and thinking about their competitive advantage, revenue streams, and business logic. These product/business insights not only enabled me to ask smart questions during interviews, but also carried me through the case/product interview sessions.

3. You should build relationships with alums. One of the biggest disadvantages for aspiring Duke data science students is that those companies don't come to Duke to recruit. So, while our software or consulting friends were talking to recruiters at career fairs, my DS friends and I found ourselves with nowhere to go.

Taking advice from Howie Rhee (Duke mentor), I cold emailed, called, and flew out to the Bay Area to build relationships with Duke alums working in the DS industry. Connecting over lunch, coffee, or drinks, I learned so much about the industry, life in the Bay Area, and the recruiting process from those good people (which is also why I feel the urge to give back). I eventually landed a few interviews through those internal referrals, and it was those opportunities I converted into offers. ∎

James Chenyang Wang is a data storyteller at Duke and an incoming product analyst at Thumbtack. He loves slide tacklings in soccer.

Data for Good
Has Growing Need for
PROJECT LEADERS

2018 was a banner year for Data for Good, full of important projects and new opportunities. The D4G community grew larger, stronger, and more connected. It also started developing growing pains.

As the concept of Data for Good has gone mainstream, the number of projects has grown dramatically—but the number of people looking to become connected with a project has grown even faster! Also, many projects today are larger, with hundreds or even thousands of participants in some of the larger civic hackathons.

The good news is there are now legions of people wanting to use their analytic skills on projects that make a real difference in their community and our world. The challenge today is capturing this lightning in a bottle and directing it and our expertise to make the most impact.

One of our greatest challenges is creating the projects, themselves, and helping people plug into them. This need highlights a growing problem of leadership in Data for Good—not in direction, but in the skills needed to design, set up, and lead the projects. Developing well-trained analytic project managers is just as critical to making an impact in Data for Good as learning the statistical skills needed to do the analysis.

Hackathons are growing in popularity, and for good reason: They usually have the following characteristics of projects that pull in volunteers:

- **Time-bound:** people need to know when it will start and end

- **Expectations:** people need to know specifically what they will be asked to do

- **Skill-led:** people need to know their particular skill set is needed and will be useful

- **Actionable:** people need to know they can make a direct impact without a lot of extra training in advance

Hackathons can also be a great place to learn the leadership skills needed for D4G projects. One way to start is by networking with the people running the event. After participating in a few, ask if you can participate on the team for the next one. There are a number of good resources for running hackathons available online, but it's best to combine them with hands-on experience.

As people gradually take on larger roles and projects, many leaders may struggle to overcome the jump from a small team personally led by the organization's head to a larger, more corporate-sized organization. Direct, personal leadership only works up to about a dozen people, which limits the tasks that can be addressed. Leaders in these organizations can move forward by collaborating on a large project involving dozens or even hundreds of people and by taking ownership of just one piece of a large project—one data source, one type of modeling method, one part of a multi-faceted problem (e.g., a survey in support of a larger effort). In the process, developing leaders learn about corporate-sized projects and develop the skills needed to collaborate with other teams on larger initiatives.

Getting Involved

Code for America's (COA) National Day of Civic Hacking is September 19. There will be multiple events in cities across the country. The emphasis tends to be coding and apps, not statistical analysis *specifically*, but analytic projects can be included. With six months to plan, it offers a good opportunity to create a project, network with others working in your area, and make a real impact. Learn more at *https://bit.ly/2QOpJtF*.

The University of British Columbia's Data Science for Social Good summer fellowship program is accepting applications through February 22. Both graduate and undergraduate students are welcome to apply. Learn more at *https://bit.ly/2RNqIUP*.

By this point, I expect a number of readers are thinking, "But we've got a great project! Where are all these people who want to get involved?" This actually points to another area of project leadership: promoting your work. While having great ideas is a critical first step, developing an understanding of communication, advertising, networking, and recruiting will go a long way toward finding the people needed to make a project successful. Website design and development, working with the media, and recruiting sponsors are important skills leaders need.

Our statistics defines a place in the "For Good" space, but organizational, project management, and leadership skills are also needed to enable more projects and help the growing number of volunteers get involved in well-organized projects that make a difference. Please send me your thoughts about this and other concerns you think the Data for Good community should address.

I hope to be able to talk with many of you February 14–16 in New Orleans at the Conference on Statistical Practice, which will again feature a Data for Good networking dinner out and a panel discussion about ethics, including the role of Data for Good. ∎

With a PhD in statistical astrophysics, **David Corliss** leads a data science team at Fiat Chrysler. He serves on the steering committee for the Conference on Statistical Practice and is the founder of Peace-Work, a volunteer cooperative of statisticians and data scientists providing analytic support for charitable groups and applying statistical methods in issue-driven advocacy.

MENTORING AWARD

ASA Mentoring Award Nominations Due March 1

Is there someone in your career history who has been a significant and ongoing mentor to you?

Has this person provided the same kind of support, guidance, and encouragement to others?

If so, consider nominating them for the ASA's annual Mentoring Award.

Established in 2015, the award is bestowed upon an ASA member who has "demonstrated extraordinary leadership in developing the careers of statistics students, statisticians, or statistical researchers early in their careers."

The award will be presented at the Joint Statistical Meetings this summer.

Send a letter of nomination outlining the reasons why the nominee deserves to be selected to *awards@amstat.org*. Additional supporting letters are encouraged.

www.amstat.org/ASA/Your-Career/Awards/ASA-Mentoring-Award

PASTIMES OF STATISTICIANS

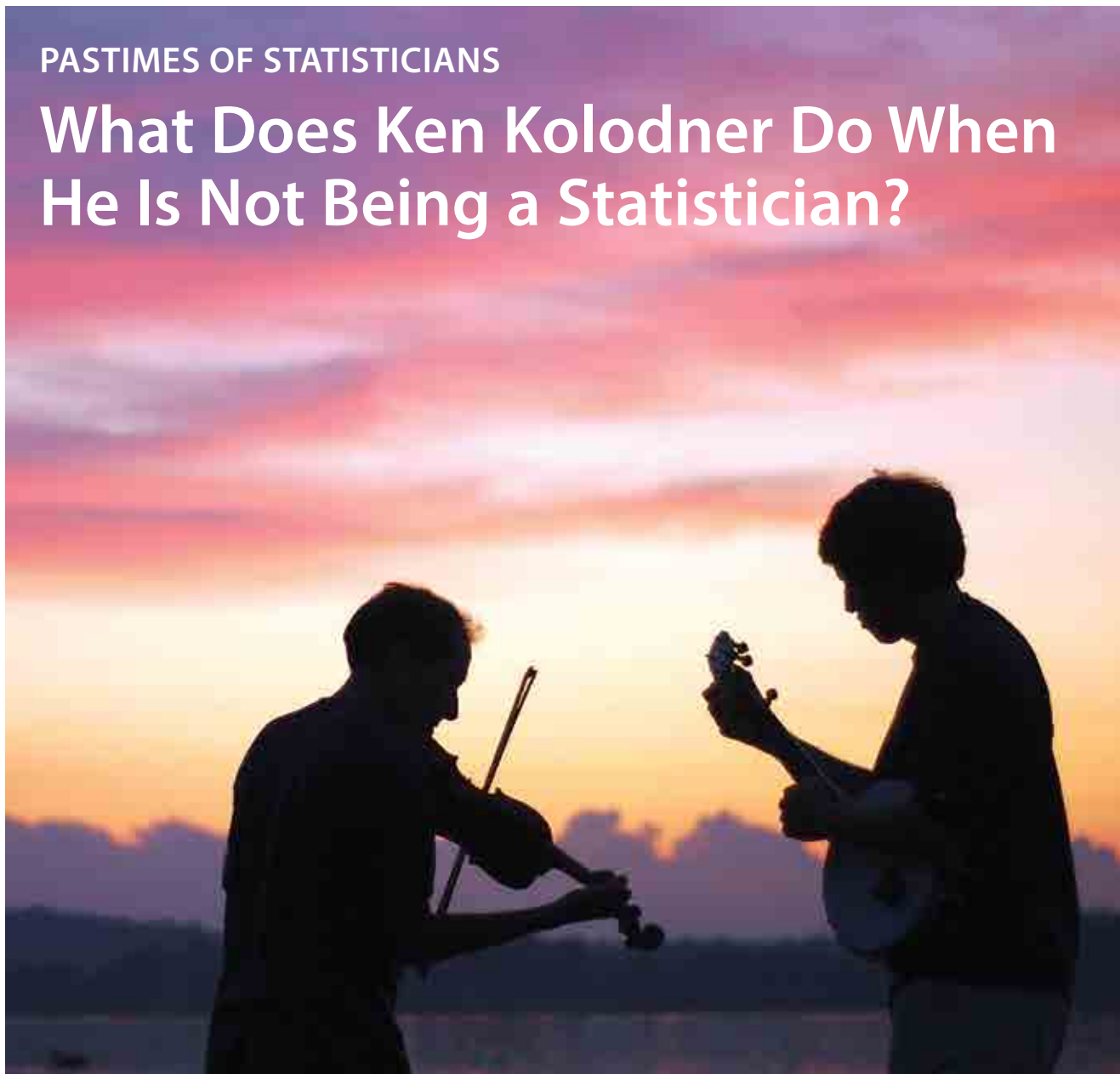# What Does Ken Kolodner Do When He Is Not Being a Statistician?

Photo Courtesy of Ken Kolodner
Ken Kolodner, left, and his son, Brad, perform at Caspian Lake in Vermont.

Epidemiologist/statistician plays music with his son

In 1978, I began working on a doctorate at The Johns Hopkins University School of Hygiene and Public Health, as it was called back then. Just prior to that and just for fun, I thought I would teach myself to play the fiddle. I had absolutely zero music background, but I thought learning to play the fiddle would be a great diversion. A few years later, I discovered the hammered dulcimer and was totally hooked.

I still looked at all this music-making as just a fun thing, not a professional endeavor. But by the time I finished my doctorate in 1985, I was performing regularly and decided I would split my time between working as a musician and an epidemiologist/statistician consultant.

Over the years, music gradually became much more of my focus. I toured extensively with a world music trio called Helicon (up to 150 concerts per year) and later with partners including Scottish National Fiddle Champion Elke Baker and Quebec fiddler Laura Risk. I performed in nearly every state in the US, recorded more than 20 CDs, published five books about playing the hammered dulcimer

Photo Courtesy of Ken Kolodner
Ken Kolodner, right, and his son, Brad, recently released a third CD.

and fiddle, and founded a four-week series of elite master classes to teach the hammered dulcimer.

As much as I love performing, I also enjoyed the challenges of doing research. There was always significant down time while touring, which gave me time to work part-time as a consultant. And in between tours, I usually found chunks of time to do consulting. Over the years, I published more than 100 articles in medical journals.

For the last five years, I have been doing music almost exclusively, largely because I have a new partner. My son, Brad, started playing the banjo at age 17 and has become one of the finest "clawhammer" banjo players in the US. He is also a fiddler like his dad.
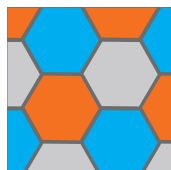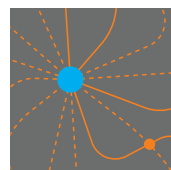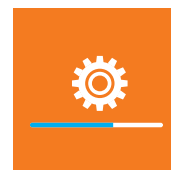
I've had many great music partners over the years, but nothing could top playing professionally with my son. I decided it was time to enjoy performing with him for as long as possible. I also perform with various other younger players. It is a blast! We just released our third CD, "The Swift House."

On other fronts, I am also totally addicted to playing soccer. I am now 62, but still play soccer 3–4 days per week and squeeze in a game of squash now and again. I can't say I have ever been bored. Life sure is busy! ∎

Recharge Your Data Work at SDSS 2019
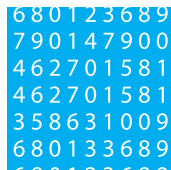
| Education | Data Science Technologies | Machine Learning | Computational Statistics | Data Visualization | Practice and Applications | Software |

While there are many aspects of conferences I enjoy, one of my favorites is when I am sitting in a talk and suddenly the speaker says something that provides unique insight into one of my own research problems. Whether it's a suggested reference, a method used, or the framing of the problem, something clicks into place and my own research agenda gets a recharge.

If your data work could use a fresh perspective, then I recommend you consider attending the 2019 Symposium on Data Science and Statistics (SDSS) in Bellevue, Washington, May 29 to June 1. Drawing together a diverse crowd of data experts, SDSS 2019 will allow its participants to discover interesting and cutting-edge techniques for extracting knowledge from data. The program committee has been hard at work putting together a program centered on the following topics:

- Computational Statistics
- Data Science Technologies
- Data Visualization
- Education
- Machine Learning
- Practice and Applications
- Software

The speakers are data experts from academia, industry, government, and nonprofits and carry titles such as statistician, data scientist, applied mathematician, computer scientist, and software engineer. Consider participating in a short course and maximizing your number of "Ah-Ha!" moments by attending as many technical talks as you can.

And, if you are interested in presenting at SDSS 2019, it is not too late to submit an e-poster abstract. E-poster abstracts are due by March 12. Poster presentations cover another one of my

SDSS
SYMPOSIUM ON
DATA SCIENCE & STATISTICS
BEYOND BIG DATA: BUILDING DATA TOOLS
BELLEVUE, WASHINGTON • MAY 29–JUNE 1, 2019

Do More

Check Out the Program *https://bit.ly/2VV5PB1*

Read About the Conference *ww2.amstat.org/sdss*

Submit an E-Poster Abstract *https://bit.ly/2SVo5Zf*

Kelly McConville is an assistant professor of statistics at Reed College and program chair for 2019 Symposium on Data Science and Statistics. She is a survey statistician who collaborates with the US Forest Inventory and Analysis Program and the US Bureau of Labor Statistics. At Reed, she teaches a variety of statistics and data science courses.

favorite conference activities: networking. Poster presentations give you the opportunity to interact directly with the speaker, and these interactions can even lead to exciting new collaborations.

Speaking of networking, we have placed many opportunities to make new connections into the SDSS 2019 program. These include a speed mentoring session, a data hackathon, themed lunch meet-ups, and a banquet.

I do hope you will consider attending SDSS 2019. You can follow the SDSS Program Committee on Twitter at @SDSSconf and tweet about the conference with #SDSS2019. And if you do attend, make sure to let me know if you have any data eureka moments! ∎

Submit a LATE-BREAKING SESSION PROPOSAL *for* JSM 2019 !

Richard Levine, JSM 2019 Program Chair

Levine

The year 2019 has arrived and JSM Denver is in sight! JSM Program Committee members began scientific program planning back in July 2018, with the final sessions submitted and scheduled in February. Check out the online program (*https:// bit.ly/2VQtuTe*) to see what we have in store for you.

Though such a lead-time is absolutely necessary for a conference this size, we realize statistical events worthy of presenting about at JSM may occur throughout the year. We thus have the late-breaking session competition, which will add up to two sessions to the JSM program, subject to approval by the Committee on Meetings.

Eric Sampson/ASA

One of the late-breaking sessions at JSM 2018, Addressing Sexual Misconduct in the Statistics Community, featured, from left to right, Keegan Korthauer, Dianne Cook, Kerrie Mengersen, Leslie McClure, Lance Waller, Kristian Lum and Brian Millen.

A late-breaking session must cover one or more technical, scientific, or policy-related topics that have arisen during the one-year period prior to JSM. These are hot statistical issues of the day and/or pressing contemporary issues in statistics.

The competition is open to any member or organization of a member society. Proposals will be judged on statistical and scientific quality, timeliness, significance and impact, potential audience appeal, and completeness.

Submitting a Proposal

Submit late-breaking session proposals to the JSM 2019 program chair, Rich Levine, at *rlevine@sdsu. edu*, with a copy to the ASA meetings department at *meetings@amstat.org* by April 15. The proposal must include the following:

- Session description, including title, summary of statistical and scientific content, explanation of the subject's timeliness and significance, and comments about the intended target audience

- Format of the session (e.g., a chair and four panelists, 2–3 speakers and a discussant, etc.)

- Names, affiliations, and contact information for the session organizer, chair, and all participants (speakers, panelists, discussants as appropriate)

- A title for each presentation in the session

- Links to relevant technical reports or news reports, if applicable

Organizers are expected to ensure participants agree to participate before the proposal is submitted. Note that a late-breaking session does not count against the JSM "one main presentation rule" (see *https://bit.ly/2A2YAhG*).

I thank every participant ahead of time for contributing to an excellent JSM 2019 program. We look forward to this last phase of program planning and then to #MakingAnImpact in Denver this summer! ∎

Recent Late-Breaking Sessions

You can find details in the programs still posted online at the ASA JSM websites (*https://bit.ly/2AHGYHW*).

2018
- Addressing Sexual Misconduct in the Statistics Community
- Statistical Issues in Application of Machine Learning to High-Stakes Decisions

2017
- National Governments, Coerced Narratives, Creative Language, and Alternative Facts
- Hindsight Is 20/20 and for 2020: Lessons from 2016 Elections

2016
- Invest in What Works: First Steps Toward Establishing Evidence-Based Policymaking Clearinghouse
- Data Journalism and Statistical Expertise: An Urgent Need for Writers, Bloggers, and Journalists to Be Statistically Savvy

2015
- The VA Secretary Bans a Statistics Book
- Meeting the Challenges of a Pandemic: The Statistical Aspects of Dealing with Ebola

2014
- Statistical Science and the President's BRAIN Initiative
- Recent Concerns About Reproducibility and Replicability: The Statistical Aspects

Teaching Statistics Trust Lecture Recording Available Online

The Teaching Statistics Trust Lecture is an annual series, organized by the Teaching Statistics Trust (TST; *www.teachingstatistics trust.org.uk*), aimed at teachers of introductory statistics students—typically in secondary schools, colleges, or early years of university—across all disciplines.

Franklin discussed her development in the teaching and learning of statistics, highlighted by what she referred to as "light bulb" moments. Topics she touched on included the following:

- The influence of key people and publications

- The use of the statistical problem-solving paradigm in teaching and learning

- Pivotal events through experience with solving personal problems using statistics

- The writing of a textbook focusing on statistics as an art and science

- The teaching of students to ask the right questions up-front (to help solve problems)

Christine Franklin, the ASA's K–12 statistical ambassador, presented the 2018 Teaching Statistics Trust Lecture, titled "Statistical Problem Solving: The Art and Science of Learning and Teaching from Data," in three locations in the United Kingdom in September of 2018.

- The co-authoring of a manual about evidence-based ways for teaching and assessing statistics at all school levels

- The chairing of a committee writing a report about the statistical education of teachers

- The recognition of the importance of producing more people educated in statistics

- The link with the emerging field of data science and problem solving using big data ∎

Feedback Requested

STEM Education

The STEM Education Coalition would like your help providing feedback to the Department of Education regarding the Education Innovation and Research (EIR) grant program and its effects on STEM education. The coalition has compiled a survey for you to provide feedback on how well aligned the department's latest round of roughly $100 million in discretionary STEM grants is with the needs of the STEM community. The survey contains explanatory instructions and links to additional background materials on the awards. All responses will be anonymous. To take the survey, visit *https://bit.ly/2Mb7IVB*.

Education Research

The Institute of Education Sciences (IES; *https://ies.ed.gov*) funds educational research and is interested in teachers' and educators' views about educational research and how they access and use it. Your input as a mathematics and statistics educator will be combined with more than 50 other educator associations to assist IES in considering how to better use their resources and reflect upon our capacities and role in bridging the gap between research/resources and our members.

Share your insights in an anonymous 15-minute survey to inform next steps at *https://bit.ly/2Max71I*. At the end of the survey, you will have the opportunity to enter your name into a drawing for one of three iPads.

No Joke: A-μ-sing Competition Entries Due April Fools' Day

Since the use of fun in teaching statistics was featured in the September 2017 issue of *Amstat News* (*https://bit.ly/2VXVJiH*), there has been much continued growth in initiatives and resources. The field's main collection at *CAUSEweb.org* has grown to nearly 800 items, including large special subcollections of songs (about 170 items), cartoons (about 180 items), and well-annotated quotes (230).

Some of the collection's visibility and growth has been sparked by contests. For example, every month since June 2016, CAUSE Cartoon Caption contests have yielded winning cartoons blindly evaluated by a judging panel for their value in teaching statistics. And the A-μ-sing Competition offers cash prizes to students and instructors who develop original fun items aligned with an educational objective.

The current A-μ-sing contest has a deadline of April 1, and the winners' work will be highlighted at the US Conference on Teaching Statistics May 16–18. Entries can be as simple as a statistical haiku or joke or as involved as a music video.

Encouraging your students to create fun examples as a class activity is a great way to get them thinking about statistical ideas in an engaging way. In fact, many A-μ-sing Competition winners have been undergraduate or graduate students whose efforts began as a class activity, including two grand prize winners and many best-in-category and honorable mentions now included in the CAUSEweb collection.

For more information about the A-μ-sing Competition and rules, visit *https://bit.ly/2FzuuFy*. ∎

Education Resources

Upcoming Conferences

NCSM Annual Conference, April 1–3, 2019, San Diego, California
https://bit.ly/2AQuzkZ

National Council of Teachers of Mathematics (NCTM) Annual Meeting and Exposition, April 3–6, 2019, San Diego, California
Stop by the ASA booth in the exhibit hall for materials and resources.
https://bit.ly/2suPlSG

US Conference on Teaching Statistics (USCOTS), May 16–18, 2019, State College, Pennsylvania
https://bit.ly/2HdxlX3

DENVER, COLORADO
JSM2019

Joint Statistical Meetings (JSM), July 27 – August 1, 2019, Denver, Colorado
ww2.amstat.org/jsm

Meeting Within a Meeting Statistics Workshop for Mathematics and Science Teachers, July 30–31
https://bit.ly/2RxFwlZ

Beyond AP Statistics (BAPS) Workshop, July 31
https://bit.ly/2FvxVh7

International Association for Statistical Education (IASE) 11th Satellite Conference, August 13–16, Kuala Lumpur, Malaysia
https://bit.ly/2FvAlXQ

National Council of Teachers of Mathematics (NCTM) Regional Conferences

September 25–27, 2019, Boston, Massachusetts

October 2–4, 2019, Nashville, Tennessee

October 16–18, 2019, Salt Lake City, Utah

https://bit.ly/2MetAQ3

Upcoming Deadlines

2019 Data Visualization Poster Competition and Project Competition
Deadline April 1
See details about the Data Visualization Poster Competition at *https://bit.ly/2LbGlHw.*

Deadline June 1
See details about the Project Competition at *https://bit.ly/2soVNdY.*

Grants, Awards, and Scholarships from the Mathematics Education Trust
Deadlines in May of 2019
The Mathematics Education Trust (MET) provides grants, awards, and scholarships for NCTM members. Start preparing proposals now for the MET grants with a May 3 deadline. See *www.nctm.org/Grants* for details.

Resources and Opportunities

New Data Science MOOC
The Johns Hopkins Data Science Lab recently released Chromebook Data Science, 12 pay-what-you-want online data science courses. These massive open online courses (MOOCs) can be completed on the Leanpub platform by anyone with basic computer skills who can read, write, and do basic math. Because of this lower prerequisite, the program can be a useful resource in K–12 data science and data literacy education. Thanks to cloud computing, all 12 courses can be completed on any computer that has access to the internet and a web browser. Another advantage of the program is that it focuses on soft skills such as communication, presentation, and how to do well in job interviews. If you are interested in using the curriculum in your classroom or school, learn more at *https://bit.ly/2VVVQvh.*
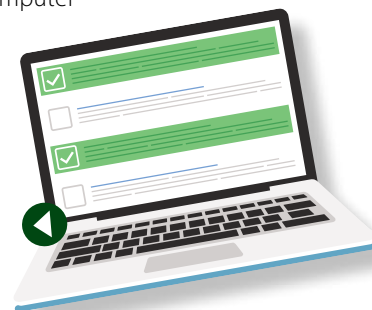
Eric Sampson/ASA
Meeting Within a Meeting takes place in conjunction with the Joint Statistical Meetings and provides K–12 teachers with hands-on activities aligned with Common Core State Standards. Attendees from 2017 listen to a session in Baltimore, Maryland.

Meeting Within a Meeting (MWM) Statistics Workshop for Middle- and High-School Mathematics and Science Teachers – Denver, Colorado – July 30–31
MWM will take place in conjunction with the Joint Statistical Meetings. The workshop is meant to strengthen K–12 mathematics and science teachers' understanding of statistics and provide them with hands-on activities aligned with the Common Core State Standards that they can use in their classrooms. The cost of the workshop is $50. More information and online registration are available at *www.amstat.org/education/mwm*. Scholarships are available.

Beyond AP Statistics (BAPS) Workshop – Denver, Colorado – July 31
The ASA/NCTM Joint Committee is pleased to sponsor a Beyond AP Statistics workshop at the annual Joint Statistical Meetings. Organized by Roxy Peck, the BAPS workshop is offered for experienced AP Statistics teachers and consists of enrichment material just beyond the basic AP syllabus. The cost of the workshop is $50. More information and online registration are available at *https://bit.ly/2DcRjx4*. Scholarships are available.

Do you know any middle- or high-school math and science teachers or experienced AP Statistics teachers who might be interested in attending these workshops? Please encourage them to attend. We are also encouraging chapters to consider sponsoring one or more teachers from their area to attend the workshops. Questions should be directed to Rebecca Nichols, the ASA's director of education, at *rebecca@amstat.org*.

2019 Data Visualization Poster Competition and Project Competition
Introduce your K–12 students to statistics through the data visualization poster and statistics project competitions, directed by the ASA/NCTM Joint Committee on Curriculum in Statistics and

Probability. There is no cost to enter either competition.

Data visualization posters (grades K–12) are due every year on April 1. See details at *https://bit.ly/2LbGlHw.*

Statistics projects (grades 7–12) are due June 1. Look closely at the new rules for the statistics project competition and a new rubric for the data visualization poster competition. See details at *https://bit.ly/2soVNdY.*

Also, judges for the 2019 statistics project competition are needed this summer. Judging takes place via email and requires about four hours of your time. If interested, email head judge, Nathan Kidwell, at *nathan.kidwell@gmail.com.*

2018 Data Visualization Poster Competition and Project Competition Winners Announced

Winners of the 2018 poster and project competitions were announced in *Amstat News* and recognized with plaques, cash prizes, certificates, and calculators. See *https://bit.ly/2sxnhOB.*

Statistics Education Webinars

The ASA offers free webinars on K–12 statistics education topics. Recently posted webinars include "Making Data Moves Using Free, Online CODAP Software with Census at School Data" by Bill Finzer, "Data in Our Daily Lives" by Anna Bargagliotti, "Building Statistical Literacy with *New York Times* Graphs" by Sharon Hessney, and "Perspectives on Data Science Education at the School Level" by Bill Finzer and Tim Erickson. This series was developed as part of the follow-up activities to the Meeting Within a Meeting Statistics Workshop. To view the webinars, visit *https://bit.ly/2QPLb1j.* NCTM also offers webinars on math and statistics education topics at *https://bit.ly/2QSAj2R*.

Build Quantitative Literacy with Graphs

What's Going On in This Graph? is a free, weekly online feature of the ASA and *New York Times* Learning Network. *New York Times* graphs of different types and context act as a springboard for middle- and high-school students in any course to think critically about graphs. Graphs are released on Thursdays and, on most Wednesdays September to April, they are discussed online. Students respond to the following three questions:

• What do you notice?

• What do you wonder?

• What's going on in this graph?

Teachers moderate in real time 9 a.m. – 2 p.m. ET on Wednesdays. At week's end, the original article, additional questions, and "stat nuggets" are revealed. No statistics background is necessary. Details are available on the ASA website at *https://bit.ly/2FvO72c.*

US Census at School Clean and Messy Data Available

The ASA's US Census at School program is a free, international classroom project that engages students in grades 4–12 in statistical problem solving using their own real data. The students complete an online survey, analyze their class census results, and compare their class with random samples of students in the United States and other participating countries. See *https://bit.ly/2GQ27jg.*

Census at School New Zealand hosts the random sampler for the international Census at School data, New Zealand data, and cleaned US data. Their online random sampler allows students and teachers to take random samples up to size 1,000 from the international, New Zealand, or US database and either download the data or start the online iNZight Lite software with the data already loaded and ready for analysis.

US Census at School provides a random sampler of USA Census at School data for download (raw, messy data).

Watch a webinar about using free online CODAP software to analyze Census at School data at *https://bit.ly/2QPLb1j.*

GAISE Now in Spanish!

Free downloads of the Spanish translation of the K–12 report and the English K–12 and college reports are available at *www.amstat.org/education/gaise.*

NCTM's *Catalyzing Change*

Today's students face a future in which there is an increasing need for mathematical and statistical skills to be used in the workplace and to make sense of our world. As a high-school teacher, leader, administrator, or counselor, part of your profession involves helping ensure students are prepared for personal and professional success. NCTM's new publication, *Catalyzing Change in High School Mathematics: Initiating Critical Conversations,* is a must-read for anyone involved in high-school mathematics education. Learn more at *www.nctm.org/catalyzing.* ∎

Emery N. Brown

Carnegie Mellon University (CMU) announced December 5 that **Emery N. Brown**—Warren M. Zapol Professor of Anaesthesia at Harvard Medical School and anesthesiologist at Massachusetts General Hospital (MGH), associate director of the Institute for Medical Engineering and Science, and Edward Hood Taplin Professor of Medical Engineering and Computational Neuroscience at the Massachusetts Institute of Technology (MIT), won the 2018 Dickson Prize in Science.

According to CMU's website, "Dr. Brown's outstanding achievements have earned him the distinction of being one of only 21 people elected to all three branches of the National Academies of Science. He is considered the 'world's expert on statistical analysis of neuronal data,' according to CMU faculty member Robert E. Kass, and his research on anesthesia has been 'truly transformative' to that field."

Brown directs an interdisciplinary team comprised of anesthesiologists, neuroscientists, bioengineers, mathematicians, neurologists, and a neurosurgeon from MGH, MIT, and Boston University (BU) that is deciphering the neuroscience of general anesthesia, CMU's announcement noted. Brown also directs the Neuroscience Statistics Research Laboratory at MGH and MIT, where the research develops statistical methods and signal processing algorithms to analyze data collected in neuroscience experiments.

Carnegie Mellon's Dickson Prize in Science is awarded annually to the person who has been judged by the university to have made the most progress in the scientific field in the United States for the year in question. At CMU, the field of science is interpreted to include the natural sciences, engineering, computer science, or mathematics.

In accepting the award, Brown credited the many people he has worked with at MGH, MIT, BU, and Harvard. "I am extremely honored to receive the 2018 Dickson Prize in Science and to join the esteemed ranks of its past recipients," he said. "I am especially grateful to all of the many students, post-docs, and colleagues whose successful collaborations have led to this recognition." To learn more about the Dickson Prize, visit *https://bit.ly/2RK4KNl.* ∎
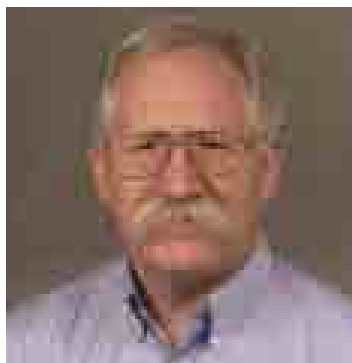
Luke Tierney

The Statistical Computing and Graphics Award committee recently chose **Luke Tierney** of the University of Iowa as the recipient of the 2019 Statistical Computing and Graphics Award in recognition of his enormously influential creation of XLisp-Stat, substantial work on MCMC methods, and critical contributions to R. A special invited panel session will be held at JSM 2019 in Denver featuring panelists who are familiar with or have been influenced by his work.

Tierney, fellow of the ASA and Institute of Mathematical Statistics, is the Ralph E. Wareham Professor of Mathematical Sciences in the department of statistics and actuarial science at the University of Iowa.

He earned his PhD in operations research from Cornell in 1980.

After being on the statistics faculty at Carnegie Mellon University and then the University of Minnesota, he joined the University of Iowa in 2002, where he chaired the department from 2004–2014. He also served as editor of the *Journal of Computational and Graphical Statistics* from 2004–2006.

Luke single-handedly built XLisp-Stat, an interactive statistical computing environment on top of the XLisp language, with many advanced features that influenced later statistical software. The newer generation of statisticians and data scientists has benefited from his fundamental contributions to R in areas such as memory management, namespacing, byte-code compiling, parallel computing, and, more recently, alternative representation for R objects. Many of these contributions are critical yet the least known to general users.

His quiet contributions have enabled a generation of data analysts to do their work through the global open source data analysis system R, which bridges statistics and data science. ∎

The Immuno-Oncology Translational Network (IOTN) Data Management and Resource-Sharing Center (DMRC) has been awarded to Roswell Park Comprehensive Cancer Center, which will serve as the overall coordination, administration, data sharing, data science, and state-of-the-art technology hub of the IOTN. **Alan Hutson**, chair of biostatistics and bioinformatics at Roswell Park Comprehensive Cancer Center and ASA Fellow, is the lead principal investigator for this project.

In addition, Hutson has been named chair of the IOTN Steering Committee.

The award of $6.28 million is in the form of a cooperative agreement between Roswell Park and the National Cancer Institute.

The DMRC will be a major national resource, providing a wealth of technical and logistical support and resources to 13 other high-profile Cancer Moonshot sites within the IOTN network. Moreover, it is anticipated that the IOTN will be expanded during the next fiscal year.

The Roswell Park investigators were selected because of their unique and wide-ranging complementary sets of expertise and technological prowess, from biostatistics, bioinformatics, immunotherapy, and immuno-oncology to Bioconductor, ontology, data science, and systems architecture. ∎

Three Win CSP Student Travel Awards

The following student winners will receive registration and travel support to attend CSP 2019.

Serifat Adedamola Folorunso, winner of the Lester R. Curtin Award, is a PhD student in the department of statistics at the University of Ibadan, Nigeria. She is a graduate statistician of the Royal Statistical Society and a member of the International Biometry Society, International Statistical Institute, Nigerian Statistical Association, Nigerian Statistical Society, and America Statistical Association.

Folorunso is currently a tutorial assistant in her department, as well as a course facilitator under the university distance learning program.

Sana Charania, winner of the John J. Bartko Award, grew up in Texas and moved to Georgia for her undergraduate education in anthropology and global health at Emory University. During her undergraduate career, she was an intern at the Centers for Disease Control and Prevention (CDC) in surveillance of parasitic diseases, which influenced her decision to pursue epidemiology at Emory's Rollins School of Public Health.

Charania currently serves as a statistics fellow on the Child Development Studies team at the CDC, where she conducts analyses using indicators of children's mental health from large population-based surveys.

She is passionate about inspiring other students to pursue a career in epidemiology or statistics and spends her free time as a mentor for epidemiology students at

Serifat Adedamola Folorunso   Sana Charania   Michelle Livings
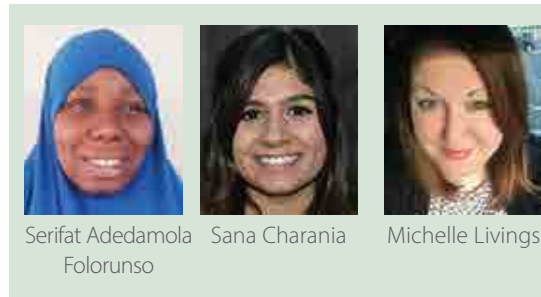
Emory University, as a SAS tutor, and as a member at large for the CDC employee association of Emory Alumni. Her favorite probability distribution is Poisson.

Michelle Livings, winner of the Lingzi Lu Memorial Award, is the surveillance biostatistician at the Southern Nevada Health District. She works primarily in the Office of Epidemiology and Disease Surveillance, compiling reportable disease statistics for Clark County, Nevada, and refining a moving average control process to detect infectious disease outbreaks. She additionally provides statistical support to other offices within the health district and serves as the statistical consultant to the Data & Systems Improvement Working Group of the Southern Nevada Homeless Continuum of Care.

Livings is also an adjunct professor of psychology at Nevada State College, teaching introductory statistics and research methods. She graduated from Georgia State University School of Public Health in December 2017 with a Master of Public Health (biostatistics concentration) degree. Her research interests include multilevel modeling with health and education data, probability sampling methods, and disease modeling and forecasting. ∎

MORE ONLINE
For information about or to register for the Conference on Statistical Practice, visit *ww2. amstat.org/csp*. For information about the travel awards, visit *https://bit. ly/2QTVwt4*.

Excellence in Statistical Reporting Award

The ASA is seeking nominations for its annual Excellence in Statistical Reporting Award (ESRA), which recognizes members of the communications media who have best displayed an informed interest in the science of statistics and its role in public life. The award can be given for a single statistical article or for a body of work.

Nominations for the 2019 award are due by March 1, and the winner will be announced in August.

In selecting the recipient, consideration is given to the following:

- Correctness, clarity, fairness, brevity, and professionalism of the communication

- Importance, relevance, and overall effectiveness in moving the intended audience

- Impact on the growth and national or regional exposure of statistics

- Appreciation and emphasis of the statistical aspects of a particular issue or event

- Excellence in coverage of research on statistics or statistical issues

Previous Winners

Kelly Servick, *Science Magazine*, 2018

Natalie Wolchover, *Quanta Magazine*, 2016

Julie Rehmeyer, *Discover Magazine*, 2015

Regina Nuzzo, Gallaudet University, 2014

Alan Schwarz, *The New York Times*, 2013

Amanda Cox, *The New York Times*, 2012

Felix Salmon, 2010

Sharon Begley, *Newsweek*, 2009

Mark Buchanan, *The New York Times*, 2008

John Berry, *Bloomberg News*, 2005

Gina Kolata, *The New York Times*, 2004

The award is a traveling cup that resides with the winner until the next winner is selected. If the winner is unable to take responsibility for the cup, the ASA will display the award at its Alexandria office.

Submitting a Nomination

Nominations may be made by editors, writers themselves, fellow writers, ASA members, statisticians, and scientists. A nomination form, nominating letter, and supporting documents are required for each nominee. The nominating letter should provide a detailed statement of the nominee's contributions to outstanding, innovative, and influential communications on statistics and the reasons for the nomination. Include copies (articles, books, videos) of the contributions of the nominee that serve as the basis of the nomination.

The nomination form may be downloaded from *https://bit.ly/2eb6Wvh*.

All documents should be mailed to Awards Nominations, American Statistical Association, 732 N. Washington St., Alexandria, VA. ∎

Ellis R. Ott Scholarship

The Statistics Division of the American Society for Quality (ASQ) awarded the following Ellis R. Ott Scholarships for Applied Statistics and Quality Management for the 2018–2019 academic year:

- **Victoria Kennerley,** University of Florida (MS category)

- **Jill Lundell,** Utah State University (PhD category)

ASQ provides $7,500 scholarships to support students who are enrolled in, or accepted into enrollment in, a master's degree or higher program with a concentration in applied statistics and/or quality management.

This includes the theory and application of statistical inference, statistical decision-making, experimental design, analysis and interpretation of data, statistical process control, quality control, quality assurance, quality improvement, quality management, and related fields. The emphasis is on applications, and studies must take place at US or Canadian institutions; online programs are excluded.

During the last 21 years, scholarships totaling more

Ott Scholarship Governing Board

Lynne Hare

J. Stuart Hunter

Tom Murphy

Dean V. Neubauer

Robert Perry

Susan Schall

Ronald Snee

J. Richard Trout

Neil Ullman

Statistics Division
ASQ The Global Voice of Quality®

than $315,000 have been awarded to 56 students.

Forms for the 2019–2020 academic year will be accepted until April 1.

Qualified applicants must have graduated in good academic standing in any field of undergraduate study. Scholarship awards are based on demonstrated ability, academic achievement, industrial and teaching experience, involvement in student or professional organizations, faculty recommendations, and career objectives.

Application instructions and forms can be downloaded at *https://bit.ly/2sw1A1b*.

For more information, contact Lynne B. Hare at *lynne.hare@comcast*.net or 55 Buckskin Path, Plymouth, MA 02360. ∎

Joint Policy Board for Mathematics Communications Award

The Joint Policy Board for Mathematics (JPBM) Communications Award is given each year to reward and encourage communicators who, on a sustained basis, bring mathematical ideas and information to nonmathematical audiences.

Up to two awards of $2,000 are made annually. Both mathematicians and nonmathematicians are eligible.

Nominations, which should include biographical information and a short description of the work that is the basis for the nomination, must be submitted by March 15 to Carla D. Savage, Office of the AMS Secretary, Department of Computer Science, Box 8206, North Carolina State University, Raleigh, NC 27695-8206.

Visit the JPBM website at *https://bit.ly/1RQEl5l* for details. ∎

How Can We Help?

We want to help you share your own news with colleagues and showcase your latest successes.

It is important to us that everyone knows about your research, recent awards, and promotions!

If you have any news you would like to share, email *megan@amstat.org.*

Deadlines and Contact Information for Select ASA National Awards, Special Lectureships, and COPSS Awards

Program	Deadline	Nominations	Questions
Statistics in Physical Engineering Sciences Award	Feb. 20, 2019	Ming Li, *mli@alumni.iastate.edu*	
Gertrude M. Cox Scholarship	Feb. 23, 2019	*awards@amstat.org*	Eloise E. Kaizar *ekaizar@stat.osu.edu*
Causality in Statistics Education Award	March 1, 2019	*educinfo@amstat.org*	
Edward C. Bryant Scholarship Trust Fund	March 1, 2019	*awards@amstat.org*	Dipankar Bandyopadhyay *dbandyop@vcu.edu*
Excellence in Statistical Reporting Award	March 1, 2019	*awards@amstat.org*	Sat N. Gupta *sngupta@uncg.edu*
ASA Fellows	March 1, 2019	*awards@amstat.org*	
ASA Mentoring Award	March 1, 2019	*awards@amstat.org*	Lillian S. Lin *ls.lin.mt@gmail.com*
Outstanding Statistical Application Award	March 1, 2019	*awards@amstat.org*	Jung-Ying Tzeng *jytzeng@stat.ncsu.edu*
Statistical Partnerships Among Academe, Industry, and Government (SPAIG) Award	March 1, 2019	*awards@amstat.org*	Kelly H. Zou *kelly.zou@pfizer.com*
Founders Award	March 15, 2019	*awards@amstat.org*	Lisa M. LaVange *lisalavange.asa@gmail.com*
Biopharmaceutical Section Scholarship Award	March 15, 2019	*https://community.amstat.org/biop/awards/scholarship*	
Links Lecture Award	May 1, 2019	*awards@amstat.org*	Arthur B. Kennickell *arthur.kennickell@gmail.com*
Lester R. Curtin Award	Oct. 15, 2019	*awards@amstat.org*	Ronald L. Wasserstein *ron@amstat.org*
Lingzi Lu Memorial Award	Oct. 15, 2019	*awards@amstat.org*	Victoria Sides *victoriasides16@gmail.com*
Deming Lecturer Award	Nov. 15, 2019	*awards@amstat.org*	Roger W. Hoerl *roger.hoerl@gmail.com*
Monroe. G. Sirken Award in Interdisciplinary Survey Methods Research	Oct. 15, 2019	*awards@amstat.org*	John L. Czajka *jczajka@mathematica-mpr.com*
Elizabeth L. Scott Award	Dec. 15, 2019	*community.amstat.org/copss/home*	

sectionnews

Physical and Engineering Sciences

Vaneeta Grover, Industrial Speakers Program Chair

The SPES Marquardt Memorial Speakers Program facilitates visits of experienced applied statisticians to colleges and universities to give a seminar and meet with students and professors. SPES reimburses the host institution up to $1,000 to cover the expenses of the speaker's visit. The speaker provides information to students about the following:

- What an applied statistician does

- How applied statisticians solve problems in science, engineering, technology, and business

- What nontechnical skills are required to be successful as an applied statistician

The Marquardt Industrial Speakers Program was established by SPES in the early 1990s to encourage careers in applied statistics. If you are 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 Vaneeta Grover at *vkgrover@yahoo.com*. ∎

Quality and Productivity

The 36th ASA Q&P Section's Quality and Productivity Research Conference will be hosted by American University in Washington, DC, June 11–13. The theme of this year's conference is "Data and Science Is a Winning Alliance."

The goal of the conference is to stimulate interdisciplinary research among statisticians, scientists, and engineers in quality and productivity. The conference aims to demonstrate and explore the marriage of data and science, as well as modern scientific approaches to handling big, multidimensional, and unstructured data. Its aim is to show convincingly how data and science—the two fundamental sources of knowledge—are using each other to produce new discoveries, progress, improved quality, and increased productivity.

The conference will include comprehensive discussions about cutting-edge methodologies in all aspects of data science and the current progress made in such computer-intensive fields as stream data mining, machine learning, functional data analysis, image reconstruction, and facial recognition. Participating statisticians, data scientists, quantitative analysts, and representatives of different branches of industry and government will exchange novel ideas and experiences working with modern big data to gain knowledge and apply it in numerous fields.

The conference will honor Barry Nussbaum, former chief statistician of the US Environmental Protection Agency and the 112th president of the ASA.

Conference registration is open. For an additional fee, the short course "Big Data Analytics: Dealing with Structured, Semi-Structured, and Unstructured Data" by Choudur Lakshminarayan of Teradata Corp. can be added to your registration. The course will be offered June 10.

American University will provide accommodation at a lower cost than nearby hotels for conference participants who choose to stay on campus. The conference website at *https://bit.ly/2CcOQRE* contains up-to-date information about the conference program, short course, registration, and hotel.

To encourage student participation, Mary G. and Joseph Natrella scholarships, QPRC student scholarships, and travel awards will be available. Further information will be provided via the conference website.

To submit a contributed paper, provide the title, authors, and an abstract to Michael Baron at *baron@american.edu*. The deadline for submitting an abstract is April 1. ∎

Biometrics

Edited by Maria Cuellar and Zheyu Wang, Biometrics Section Publications Officers

The Biometrics and Mental Health Statistics sections are sponsoring a webinar about smartphone-based digital phenotyping, to be presented by J.P. Onnela January 24 from 12:00 p.m. to 1:30 p.m.

Our evolving data ecosystem is creating many novel data streams for conducting health research, and the smartphone is one new means to collect data on human health and behavior. Onnela, associate professor of biostatistics at Harvard University, will discuss his experience using smartphones to define patient phenotypes and create a platform for novel data collection and study implementation.

Registration is open. Webinar details are available at *https://bit.ly/2RKkuQr*. ∎

The NC ASA community enjoys food and conversation during its fall dinner.

Stories of Significance Highlight North Carolina Chapter's Fall Dinner

Nearly 100 statisticians from the North Carolina ASA community gathered for the North Carolina Chapter's 2018 Fall Dinner November 30. The theme was "Stories of Significance." 2018 ASA President Lisa LaVange and Duke University's Mine Çetinkaya-Rundel were the speakers.

Çetinkaya-Rundel, associate professor of the practice and 2016 ASA Waller Education Award honoree, spoke about her move from actuarial science to statistics education. She discussed important realizations and people with impact throughout her educational and career journey. She also talked about the unexpected nature of her path, the importance of learning new skills, and how following her passion led to doing what she loves.

LaVange followed with her own stories of significance as she spoke about the impact of specific decisions on her career trajectory. In addition to talking through her process as she moved through her career, LaVange also said leadership is a science. Additionally, she addressed the importance of diverse role models. ∎

Duke professor Mine Çetinkaya-Rundel talks about her career path from actuarial science to statistics education.

2018 ASA President Lisa LaVange discusses her own stories of significance in a talk about how specific decisions affected her career trajectory.

The following events are the latest additions to the ASA's online calendar of events. Announcements are accepted from education and not-for-profit organizations only. To view the complete list of statistics meetings and workshops, visit *www.amstat.org/dateline*.

* Indicates events sponsored by the ASA or one of its sections, chapters, or committees

›› Indicates events posted since the previous issue

10–12—The 7th Workshop on Biostatistics and Bioinformatics, Atlanta, Georgia
For details, contact Yichuan Zhao, 1342, 25 Park Place, Atlanta, GA 30303; (404) 413-6446; *yichuan@gsu.edu*.

13–17—NSF-CBMS Regional Conference on Topological Methods in Machine Learning and Artificial Intelligence, Charleston, South Carolina
For more information, visit *https://bit.ly/2HkKSMP* or contact Ben Cox, 66 George St., Charleston, SC 29424-0001; (843)953-4973; *coxbl@cofc.edu*.

*29–6/1—2019 Symposium on Data Science & Statistics, Bellevue, Washington
For more information, visit *ww2.amstat.org/sdss* or contact ASA Meetings, 732 N. Washington St., Alexandria, VA 22314; (703) 684-1221; *meetings@amstat.org*.

2019

March

18–22—German Joint Statistical Meeting DAGStat2019, Munich, Germany
For more information, visit *www.dagstat2019.de* or contact Michael Lebacher, Ludwigstr. 33, Munich, International 80539, Germany; +49 89 2180 2226; *michael.lebacher@stat.uni-muenchen.de*.

May

2–4—SIAM International Conference on Data Mining (SDM19), Calgary, Alberta, Canada
For more information, visit *https://bit.ly/2Nmr49p* or contact Eva Donnelly, 3600 Market St., 6th Floor, Philadelphia, PA 19104; (215) 382-9800; *meetings@siam.org*.

June

››*11–13—2019 Quality and Productivity Research Conference, Washington, DC
For details, visit *https://bit.ly/2CcOQRE* or contact Emmanuel Yashchin, IBM T.J. Watson Research Center, Box 218, Yorktown Heights, NY 10598; (914) 945-1828; *yashchi@us.ibm.com*.

»12–14—BISP11 - Eleventh Workshop on Bayesian Inference in Stochastic Processes, Madrid, Spain
For more information, visit *https://bit.ly/2AG2Oel* or contact Fabrizio Ruggeri, Via Alfonso Corti 12, Milano, International 20133, Italy; +39 0223699532; *fabrizio@mi.imati.cnr.it*.

16–19—39th International Symposium on Fore-casting, Thessaloniki, Greece
For details, visit *isf.forecasters.org* or contact Pamela Stroud, 53 Tesla Ave., Medford, MA 02155; (781) 234-4077; *isf@forecasters.org*.

18–21—The 7th International Workshop in Sequential Methodologies, Binghamton, New York
For details, visit *sites.google.com/view/iwsm2019* or contact Aleksey Polunchenko, 4400 Vestal Parkway East, Binghamton, NY 13902; (607) 777-6906; *iwsm2019@gmail.com*.

Center For Clinical Epidemiology And Biostatistics 12th Annual UPENN Conference on Statistical Issues in Clinical Trials

April 17, 2019

Electronic Health Records (Ehr) In Randomized Clinical Trials: Challenges And Opportunities

Website and Registration opens January 3, 2019
https://bit.ly/2QLGLIP

CASE STUDIES

Denise Esserman, Phd, *Yale University:* From Screening to Ascertainment of the Primary Outcome Using Ehr, Challenges in The Stride Trial

Steven Zeliadt, Phd, *University Of Washington:* The Approach Trial: Assessing Pain, Patient Reported Outcomes and Complementary and Integrative Health

Richard Platt, Md, *Harvard University*: The Impact-AFib Trial: **IM**plementation Of An RCT to Im**P**rove Treatment with Oral **A**nti**C**oagulan**T**s in Patients with **A**trial **F**ibrillation

Matthew T. Roe, Md, *Duke University*: Leveraging Electronic Health Record Data forPragmatic Randomized Trials in Learning Health Care Systems in the United States – Lessons Learned From The Adaptable Trial

OVERVIEWS

Benjamin A. Goldstein, PhD, *Duke University:* Design Considerations for Running Health System Based Trials Through the Electronic Health Record

Mark LevensOn, PhD, *Food and Drug Administration:* Regulatory Perspective On Performing Rcts in an Ehr Environment

STATISTICAL METHODS

Patrick J. Heagerty, PhD, *University of Washington:* Addressing Heterogeneity inthe Data, Design, and Analysis of Pragmatic Trials Embedded in Delivery Systems

Sebastien Haneuse, PhD, *Harvard University:* On Selection Bias Due to Missing Data in EHR-Based RCTs

Susan M. Shortreed, PhD, *Kaiser Permanente Washington Health Research Institute:* Using Real-World Data to Improve Trial Design

JINBO CHEN, PhD, *University of Pennsylvania:* Study Design Issues for Exploiting EHRs to Design Clinical Trials

PANEL DISCUSSION

Rebecca Hubbard, PhD University of Pennsylvania

Ralph D'Agostino, PhD Boston University

BILL CAPRA, PhD Genentech
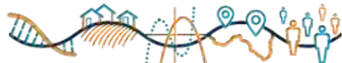
July

»3–5—ISSAT International Conference on Data Science in Business, Finance, and Industry (DSBFI 2019), Da Nang, Vietnam
For more information, visit *https://bit.ly/2VLEf9m* or contact Michelle Pham, P.O. Box 281, Edison, NJ 08818; (732) 491-5870; *michelle@issatconferences.org*.

8–12—International Workshop on Statistical Modelling 2019 (IWSM2019), Guimarães, Portugal
For details, visit *https://bit.ly/2Cixm6e* or contact Luís Machado, Department of Mathematics and Applications, University of Minho, Guimarães, International 4800-058, Portugal; 351253510443; *lmachado@math.uminho.pt*.

22–26—European Meeting of Statisticians (EMS 2019), Palermo, Italy
For more information, visit *https://bit.ly/2TDQtic* or contact Angelo Mineo, Viale delle Scienze, Ed. 13, Palermo, International 90128, Italy; *ems-2019@unipa.it*.

DENVER, COLORADO JSM2019 JULY 27–AUGUST 1

*27–8/1—2019 Joint Statistical Meetings, Denver, Colorado
For details, visit *ww2.amstat.org/jsm* or contact ASA Meetings, 732 North Washington St., Alexandria, VA 22314; (703) 684-1221; *meetings@amstat.org*.

»28–8/11—Summer Seminar in Philosophy of Statistics, Blacksburg, Virginia
For more information, visit *summerseminarphilstat.com* or contact Jean Miller, Philosophy Department, 229 Major Williams Hall (0126), Virginia Tech, Blacksburg, VA 24061; (540) 998-1123; *jemille6@vt.edu.*

August
»1–3—ISSAT International Conference on Data Science and Intelligent Systems (DSIS 2019), Las Vegas, Nevada
For more informaion, visit *https://bit.ly/2FmC6vB* or contact Michelle Pham, P.O. Box 281, Edison, NJ 08818; (732) 491-5870; *michelle@issatconferences.org.*

17–19—The Fourth Workshop on Higher-Order Asymptotics and Post-Selection Inference (WHOA-PSI), St. Louis, Missouri
For more information, visit *https://bit.ly/2NmpDYz* or contact Todd Kuffner, 1 Brookings Drive, St. Louis, MO 63130; *kuffner@wustl.edu.*

19–23—NSF-CBMS Regional Research Conference: Fitting Smooth Functions to Data, Austin, Texas
For details, contact Arie Israel, 3925 W. Braker Lane, Suite 3.340, Austin, TX 78759-5316; (512) 471-6424; *arie@math.utexas.edu.*

THE ASA BIOPHARMACEUTICAL SECTION REGULATORY-INDUSTRY STATISTICS WORKSHOP

September

*23–25—2019 ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop, Washington, DC
For more information, visit *https://bit.ly/2D3JJFd* or contact ASA Meetings, 732 North Washington St., Alexandria, VA 22314; (703) 684-1221; *meetings@amstat.org*.

October

10–12—The 3rd International Conference on Statistical Distributions and Applications (ICOSDA 2019), Grand Rapids, Michigan
For more information, visit *https://bit.ly/2OBuBFv* or contact

Felix Famoye, Department of Mathematics, Mt. Pleasant, MI 48859; (989) 774-5497; *felix.famoye@cmich.edu*

2020

January

*6–8—2020 International Conference on Health Policy Statistics (ICHPS), San Diego, California
For more information, visit *https://bit.ly/2VLmmY4* or contact ASA Meetings,

732 North Washington St., Alexandria, VA 22314; (703) 684-1221; *meetings@amstat.org*.

June

24–27—5th International Workshop on Functional and Operatorial Statistics (IWFOS 2020), Brno, Czech Republic
For details, visit *https://bit.ly/2sl8vKF* or contact David Kraus, Kotlárská 2, Brno, International 611 37, Czech Republic; *david.kraus@mail.muni.cz*.

August

»*1–6—2020 Joint Statistical Meetings, Philadelphia, Pennsylvania
For more information, contact ASA Meetings, 732 North Washington St., Alexandria, VA 22314; (703) 684-1221; *meetings@amstat.org*. ∎

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.

These listings and additional information about the 65-word ads can be found at *ww2.amstat.org/ads.*

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 *https://jobs.amstat.org/jobseekers.*

Connecticut

■ The department of mathematics at Quinnipiac University's College of Arts & Sciences invites applications for a tenure-track assistant professor position in mathematics with an emphasis in statistics or applied mathematics. For more information and to apply: *https://bit.ly/2RrlEkc.* Quinnipiac University is an equal opportunity employer.

Maryland

■ The Emmes Corporation in Rockville, MD, a full-service contract research organization, has openings for PhD-level statisticians to serve on and lead multi-disciplinary project teams supporting clinical research with great public health impact across a range of disease areas. Requirements: Solid background in statistical methods with a PhD in biostatistics/statistics/ epidemiology, strong oral and written communication skills, and leadership potential. Apply directly online at *www.emmes.com.* EOE.

Massachusetts

■ Applications are invited for a full-time visiting assistant professor in statistics with a three-year appointment to begin on July 1, 2019. Responsibilities include teaching four undergraduate statistics courses a year and helping with the comprehensive evaluation of senior majors. See the full ad and submit applications at *MathJobs.org.* All applications received by February 22, 2019, will be guaranteed consideration. Questions can be addressed to *mathstats@amherst.edu.* Amherst College is an equal opportunity employer and encourages persons of all genders, persons of color, and persons with disabilities to apply.

New York

■ The Department of Biological Statistics and Computational Biology Cornell University invites applications for a lecturer position. Qualifications: An MS in statistics or a closely related field required. A PhD is preferred. Candidates should submit a cover letter, curriculum vitae, teaching statements, statement of diversity/ equity/ inclusion and three reference letters submitted to *https://bit.ly/2TP2zoR.* Equal opportunity affirmative action educator and employer.

Texas

■ Tenure-track positions in statistics and mathematics. The University of Texas at Dallas. The department of mathematical sciences within the School of Natural Sciences and Mathematics invites applications for assistant professor tenure-track faculty in mathematics and statistics. For more information and to apply, *https://bit.ly/2Fv2off*. The University of Texas at Dallas is an equal opportunity/equal access/affirmative action employer committed to achieving a diverse and inclusive community. ■

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*.

We asked followers to tell us: What is the one character trait you need to be a statistician?

TELL US!

Can you name three women who were pioneers in the statistics profession?

Frank Harrell • **@f2harrell**
Mimic Sherlock Holmes

Mikhail Popov • bearloga
Thought about answering with "inquisitiveness" but decided I'd much rather:
- be & inspire others to be
- hire statisticians who are
- teach statistics students to be EMPATHETIC, rather than inquisitive (if I absolutely must pick one). Empathy should be at the core of what we do.

Craig A. Rolling • @rollingstats
Skepticism

Calvin Miller • @clvnmllr
Humility. A field built around uncertainties is not forgiving to those who fail to concede the possibility that they may, in fact, be wrong.

Iyue Sung • @IyueSung
Curiosity.
It begets others, like: skepticism, creativity, and persistence.

James Adams High tolerance for frustration.

Lau Lik Nang Diligent.

Xu Alex Significant

Tapio Nummi Accept uncertainty.

Ehsan Masoudi patience (with your students, non-statistician boss and clients)

Jake Humphries You can be a statistician without any of these character traits, but to be an excellent one, you'll likely need at least some of all of them.

Jana Asher some of all of them!

Iram Hasan Unbiased towards data interpretation.

Jana Asher Self-awareness. You need to understand your own biases so they don't enter into your analyses.

Statistics

The latest release of SAS/STAT® is now available. SAS/STAT 15.1 provides new methodology and new capabilities:

SAS/STAT 15.1 Highlights

Bayesian generalized linear mixed models.

Graphical causal models.

Regression for time-to-event data based on restricted mean survival time.

Counterfactual analysis using quantile regression.

Semiparametric proportional hazards model for interval-censored data.

Recent SAS/STAT Additions

Casual mediation analysis.

Compartmental models for pharmacokinetic analysis.

Fast quantile process regression.

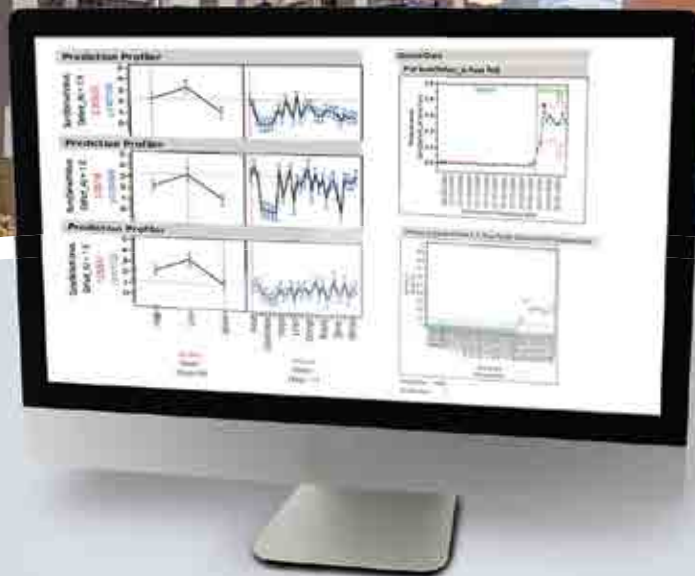Cause-specific proportional hazards analysis for competing-risks data.

Variance estimation by the bootstrap method for survey data analysis.

Learn more
support.sas.com/statnewreleases

§sas

Ed Hutchins, Cree Product Engineering Manager

Great software in the right hands can change the world.

Cree's engineers are innovating better, more efficient LED solutions for a brighter future. With analytics, brilliant things happen.

Read about Cree's success, and find out how JMP can help you change your world:

www.great.jmp

jmp®

Statistical Discovery.™ From SAS.