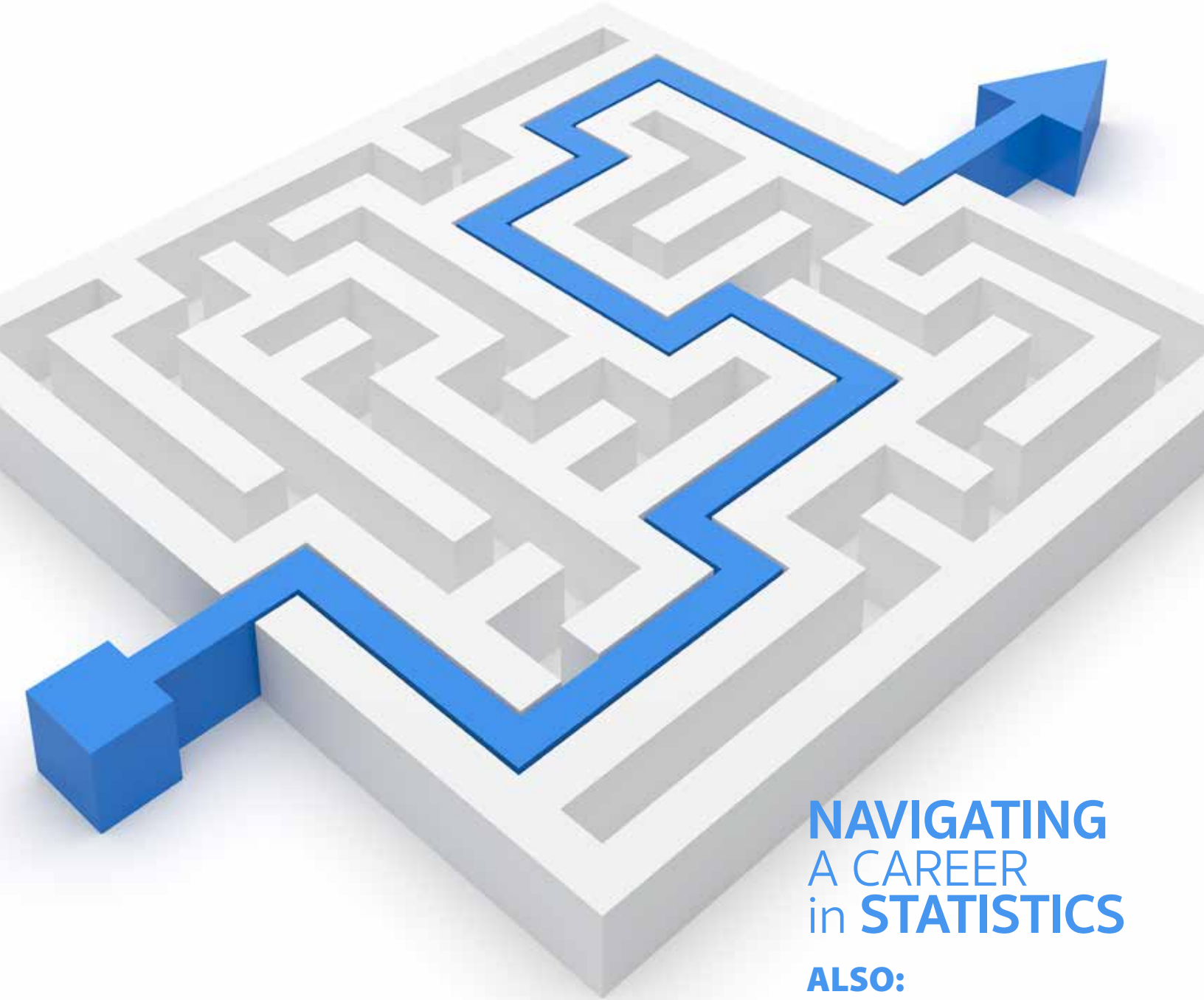
September 2018 • Issue #495

AMSTATNEWS

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



NAVIGATING A CAREER in STATISTICS

ALSO:

Rising Star Talks About Statistics Journey, Future of the Profession

Tech Leaders Discuss Working in Industry

JSM2018
Vancouver, British Columbia, Canada

Thank you 2018 SPONSORS

PLATINUM

abbvie

sas

GOLD

Lilly

FDA

IBM

MERCK
INVENTING FOR LIFE

NETFLIX

jmp
Statistical Discovery™ From SAS

Takeda

DiDi

SILVER

Bank of America

Berry Consultants
Statistical Innovation

CAMBRIDGE
UNIVERSITY PRESS

Cytel

facebook

Minitab

RTI
INTERNATIONAL

STATA

SCM
STEVENS
CAPITAL
MANAGEMENT

Taylor & Francis
Taylor & Francis Group

TWO SIGMA

Westat

XLSTAT
Your data analysis solution

AMSTATNEWS

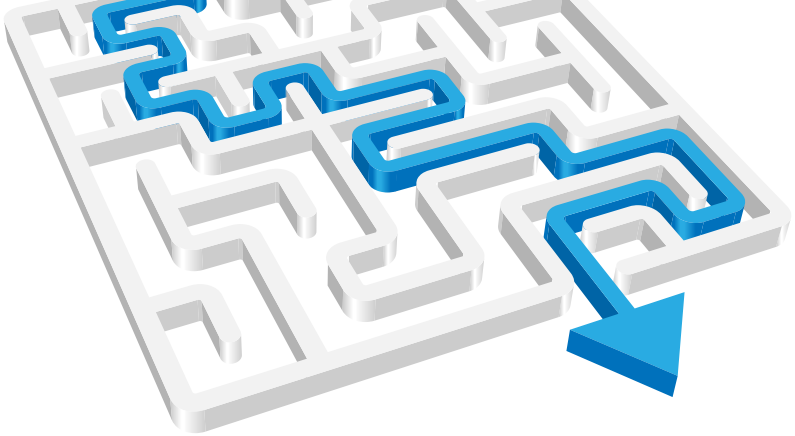
SEPTEMBER 2018 • ISSUE #495

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

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

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

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



influential statisticians

shared experiences

employer's view

career paths

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

columns

member news

Q&A

Page 38

BOB KOMARA
Data Scientist at MSX International

SCARLETT L. BELLAMY
Professor and Director of Graduate Studies in Biostatistics, Drexel University Dornsife School of Public Health

JULIE NOVAK
Senior Data Scientist, Netflix

LISA LIX
Director of the Data Science Unit in the George and Fay Yee Center for HealthCare Innovation

With a Little Help from My Friends: Advice from the Board for Your Career Journey

This month's issue of *Amstat News* focuses on careers and how the ASA can help. At JSM last week in Vancouver, I was reminded of the terrific resources provided to attendees and employers alike through the Career Service, located onsite.

Hundreds of employers and job candidates make connections each year through the Career Service. At FDA, where recruitment was a continuing need, we relied on the Career Service to identify interested and qualified candidates. If you missed JSM, you can take advantage of the ASA JobWeb (*www.amstat.org/ASA/Your-Career/Jobweb*).

Another resource for ASA members in their pursuit of career advancement is the Committee on Career Development (*http://community.amstat.org/ccd/home*). The committee's mission is to "organize and promote activities, information, and materials to aid in the career development of statisticians and data scientists and to excite new interest in careers in these fields." One resource the committee maintains is a list of professional development resources (*www.amstat.org/asa/your-career/professional-development.aspx*).

Each year, just prior to JSM, the ASA Board meets to conduct a variety of business matters. I took the occasion of this issue's focus on careers to solicit advice from members of the ASA Board: **Barry Nussbaum** (ASA past-president and retired chief statistician of the Environmental Protection Agency), **Karen Kafadar** (ASA president-elect and chair and commonwealth professor of statistics at the University of Virginia), **G. David Williamson** (ASA vice president and senior scientist at the US Centers for Disease Control and Prevention), **Kathy Monti** (ASA vice president and retired chief statistical scientist at Rho, Inc.), **Katherine Halvorsen** (ASA Council of Sections representative and professor of mathematics and statistics at Smith College), **Jim Lepkowski** (ASA Council of Sections representative and professor emeritus at the University of Michigan), **Scott Evans** (ASA publications representative and professor and director of the biostatistics center at The George Washington University), and **Ron Wasserstein** (ASA executive director). Here are some pearls of wisdom garnered from each of their distinguished and unique careers:

Q1. What is the best career advice you received?

Katherine: Persistence matters.

Scott: Think first, then research, then think again.

Kathy: Get a degree in statistics or biostatistics, rather than pure math. There are better job opportunities in stats-related fields. That was true then, and probably truer now.

Ron: Don't wait for opportunities to come along. Volunteer to take on the extra job, or work on that committee, or help with that project.

Karen: Listen carefully to the problem. Often (s)he is not asking the right question.

Barry: My supervisor asked, "Just who do you think can read this memo?" It made me realize that we must always keep the customer in mind.

David: Have a strong work ethic to always do the best you can—then your career development and progress will take care of themselves.

Jim (*from his dad, when worried about passing up a career opportunity that came at the wrong time*): "There is always another street car," referring to the Detroit of his childhood, where street cars stopped every couple of blocks and ran all the time. … Someone with as much talent and preparation and experience would always find another opportunity. … Have the confidence that opportunities would be there when you need them.

Q2. What tips can you offer for cover letters and résumés when applying for a job?

Barry: It is likely they will only read the cover letter, not the full CV. So step out fast and succinctly. In the CV itself, tell why you

Photo by Jon Gardiner/UNC-Chapel Hill

Lisa LaVange

From left: Ron Wasserstein, Lisa LaVange, and Karen Kafadar during the ASA Board Meeting at the Joint Statistical Meetings 2018

did some of your projects and give some substantive results.

Jim: A good proof reading, a few minutes' extra effort, would have gone a long way to making sure that what I knew about them through the written materials gave a good impression.

Kathy: Explain why you are interested in the job and what you might bring to the job in the cover letter. Make the CV relevant to the job and well organized. Respect the page limits imposed by the application and in no case should one pad with long lists of minutiae. Perhaps the CV for the first job might include courses taken, but later in the career, this would be considered padding.

Karen: Make it personal, and, for heaven's sake, correct spelling and grammatical errors. Padding the CV will leave a negative impression.

David: Construct your résumé so that it's individualized to showcase your strengths to match with the needs and mission of the employer to whom you're sending the resume.

Scott: Tailor the cover letter to the requirements expressed in the position announcement. For the CV, iterate several times to optimize organization, clarity, and efficiency. Keep it up to date.

Katherine: Keep the cover letter short and to the point. Have a friend or colleague review it before you send it.

Q3. What general career advice would you like to share?

Karen: Don't hesitate to ask for help when needed.

Scott: Never stop being a student.

Kathy: Your education starts in school, but it doesn't end there. Try to keep current in methods that are pertinent to your field once you move into the job market. Getting involved with the ASA's chapters and sections can be an important part of your career. When thinking about your job options, don't ignore the feel of the place.

Katherine: Don't take a job because it pays well. To make the work worth doing, you need to like what you do and feel that what you are doing is meaningful.

Ron: Don't be afraid to take risks, and don't take rejection (as in not selected for the job) personally.

Barry: John Tukey once said, "The great thing about being a statistician is that you get to play in everyone else's sandbox." I have appended this with Nussbaum's third postulate: "The person who owns the sandbox is impatiently waiting for a sturdy, useful sand castle to be constructed."

David: The ball is in your court regarding your career. Work to gain personal traits and interpersonal skills that facilitate success (e.g., initiative, strong ethics, ability to communicate effectively orally and in writing), take training to acquire strong technical and nontechnical skills, find and be a good mentor, and identify your interests and passions and lay out your career path to integrate them. Strive to make a positive impact in as much of your job (and life) undertakings as possible. Make a difference!

When asked what courses they did not take but wished they had, responses ranged from the general—take more science courses—to the specific—chemistry, public speaking, logic, and theology. Jim said, "I wish I had a good course in systematic theology to help me begin to order thinking about what really matters."

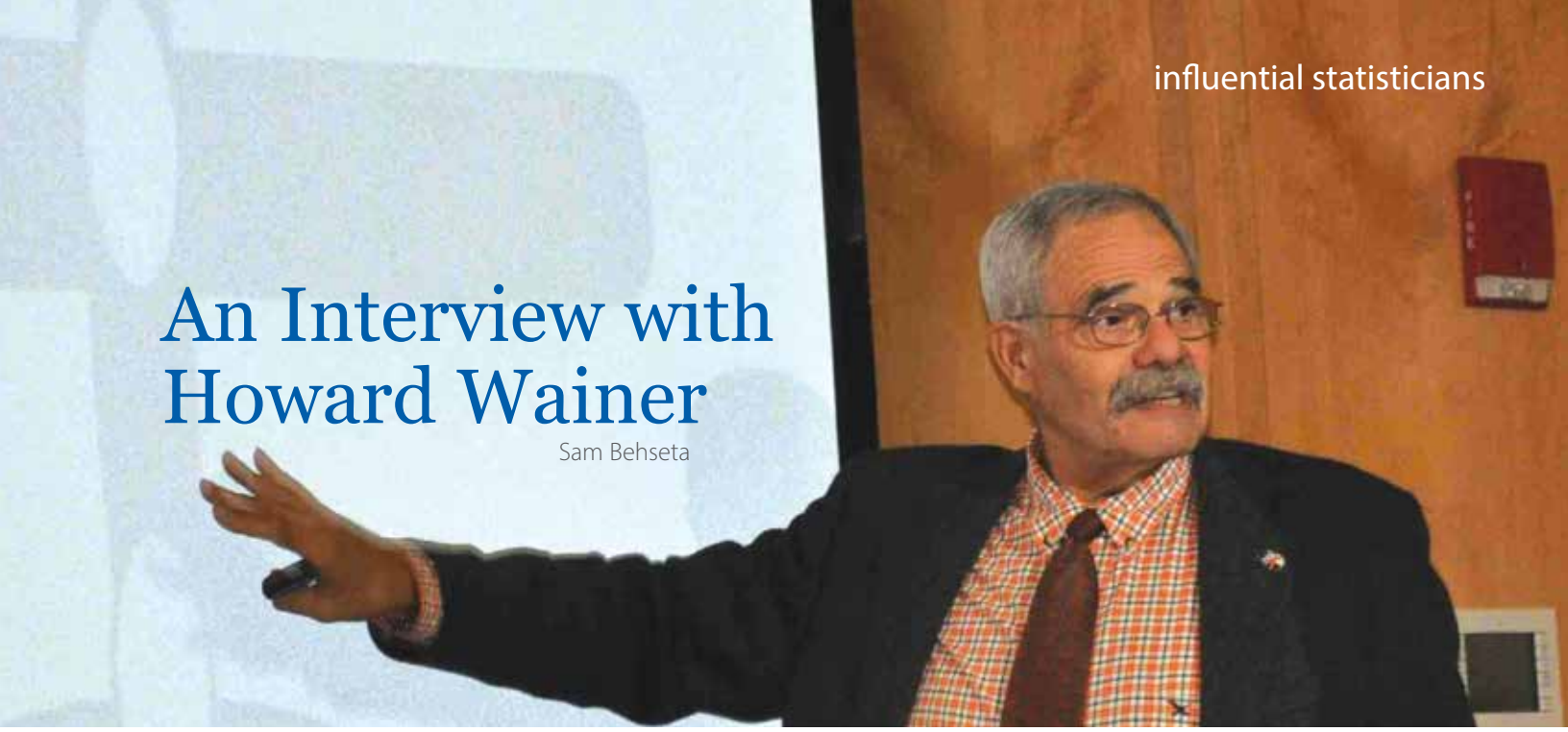
Thanks to these board members for their inspiring words of advice and encouragement. And thanks to the ASA for continuing to find ways to help members as they navigate their own career paths!

Sometimes, a second career warrants notice, and later in this issue, you can read about one such instance. Maura Stokes, senior scientist and director at SAS, published her first novel for young adults this summer, *Fadeaway* (Yellow Jacket, Bonnier Publishing USA, 2018). Whether a statistician writing on the side or vice versa, Maura has clearly managed to navigate a very successful career while still finding time for the things she is passionate about. A shout-out to Maura, whom I am proud to call a friend. And read the book—it is terrific! My grand-daughters have signed copies waiting for them (to learn to read!)

Lisa LaVange

An Interview with Howard Wainer

Sam Behseta



Howard Wainer gives a talk during his book signing (*Uneducated Guesses*) at the Princeton Library.

Howard Wainer, who writes the Visual Revelations column for *CHANCE* magazine, is distinguished research scientist (retired) at the National Board of Medical Examiners. He has won numerous awards and is a Fellow of the American Statistical Association and American Educational Research Association. His interests include the use of graphical methods for data analysis and communication, robust statistical methodology, and the development and application of generalizations of item response theory. He has published 20 books so far; his latest is *Medical Illuminations: Using Evidence, Visualization & Statistical Thinking to Improve Healthcare.*

Howard Wainer's Visual Revelations column—appearing in *CHANCE* magazine, which is published by the ASA—is arguably the longest-running serial publication in the history of statistics: 25+ years of delving into the core events of the time, all from the point of view of a razor-sharp, uncompromising, and greatly responsible statistician. Now, that is a career unto itself, but then there are some 400 articles, volumes of books, and years of leadership at the Educational Testing Service and National Board of Medical Examiners—not to mention a long teaching career at the University of Pennsylvania. In recognition of his extraordinary contributions to *CHANCE* and the profession in general, former *CHANCE* Executive Editor Sam Behseta sat down with Howard for a friendly chat.

Sam Behseta: Howard, how did you become involved with *CHANCE*?

Howard Wainer: I got a call from Steve Fienberg after he and Bill Eddy started *CHANCE*. He knew of my interest in graphics and said there's a column called Visual Revelations that Alan Paller had originated. Alan had written the first year or two's worth, but soon he was to be leaving, so Steve asked if I would take it over for a while. I thought about it. I had ideas for about three columns, and thought that should carry me through most of the first year. And, surely, I would be able to think of something else for the fourth one. So, I figured I could certainly take it for a year and, with luck, I could get through two. So, I agreed.

Every year since, I've thought, "Well I've got another idea for one more column," and so it just kept going. A bit like mathematical induction—if I had *n*-columns, I could always dream up one more. Anyway, I have enjoyed it so much that I just kept going. It has quickly been 25 years.

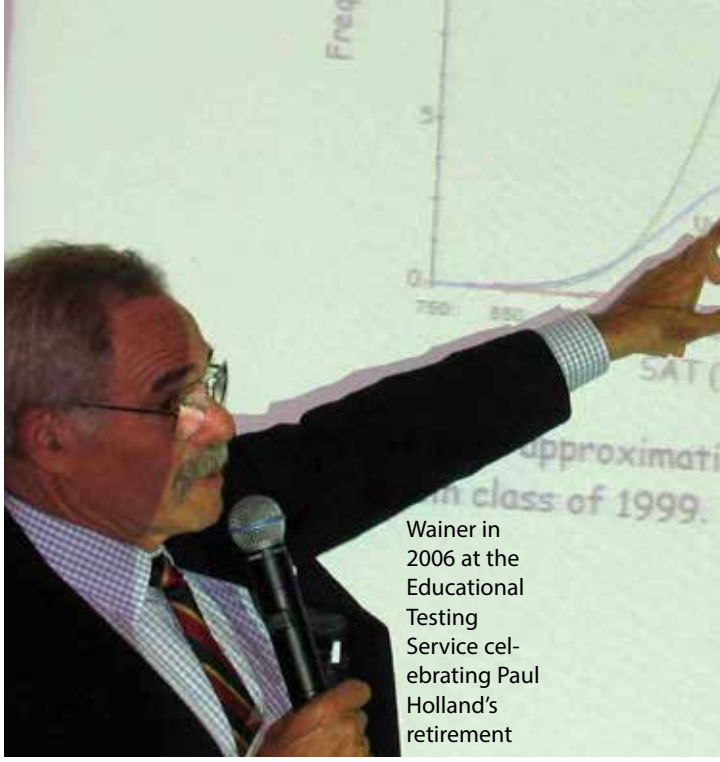
Sam Behseta: So, what's the process for you? How does this work? I mean, obviously, you look around and observe and you have your professional stuff, but then how does it work? How do you decide a piece is suitable for *CHANCE*?

Howard Wainer: To some extent, it's like relying on the kindness of strangers. I've relied on the kindness and wisdom of the editors, and the editors of *CHANCE* have been both wise and generous over the course of all these years. I figure about one idea out of every 50 is any good. Some people get one idea a week, and so they get one good one a year. John Tukey said he used to get 50 ideas before breakfast and, although even he—he claimed—had only one good one out of 50, it didn't matter because he had plenty of them.

I don't have 50 before breakfast, but I have a bunch of them. But there's a filtering process that goes on. I pick one I think will be good, and then I draft something

Sam Behseta earned his PhD from Carnegie Mellon University and is a professor at California State University, Fullerton. His main research area is statistics in neuroscience. Behseta has both trained and mentored undergraduate and graduate students.

Wainer and his wife, Linda, on Hurricane Mountain in the Adirondacks



Wainer in 2006 at the Educational Testing Service celebrating Paul Holland's retirement

and send it to the editor and ask if it's suitable? Obviously, I'm relying both on wisdom and generosity for the editor to say, "Yes, it is" or "No, it isn't."

Sam Behseta: Is there any cause célèbre you see as a source of future topics?

Howard Wainer: I'm becoming more convinced it's going to be useful to take a more publicly Bayesian point of view now. I've been reading more about what is called the prosecutor's fallacy. Basically, it's calculating the wrong probability and not knowing how to do the Bayesian flip (or not realizing it should be done). Maybe that will be worth describing.

But the image of the initial audience for my column I have in mind is often that of the Stat 101 teacher who's trying to present an idea to their class to illustrate a statistical principle. It has to be something they can do in one class without much math beyond algebra. My vision is that an instructor could go through the various columns and pick out a topic—a case study they can use to illustrate something the students can read and get some idea of how exciting it is to be in the field we are in.

Sam Behseta: What I noticed in your pieces is there is a sense of justice. You're interested in social justice, largely speaking. Where does that come from? Do you feel a responsibility to comment on social issues? Or does it just come with it?

Howard Wainer: I'm not sure. If I had to say I had a religion, it's the worship of evidence. You look at the evidence and follow where it leads. I guess that's what you might call a weak prior in some important sense. In addition, my parents were children of the Depression, and so FDR was much closer to being a deity than a politician. So, the 12th commandment, certainly in my family, was thou shalt not vote Republican. So, I more or less absorbed democratic as well as Democratic ideals as part of mother's milk.

Sam Behseta: So, you were in Chicago in the '60s? I was talking to Steve Fienberg and he was saying you were more or less there at the same time.

Howard Wainer: Yes. We came pretty much together. He was there a year ahead of me in '69, and I came in '70.

Sam Behseta: But you didn't study statistics?

Howard Wainer: I was a faculty member, not a student. I was in something called the Committee on Methodology in the department of behavioral sciences.

I met Steve early in the fall of '70; I had just arrived a few months before. They recruited two young assistant professors in statistics to count the votes in the faculty senate election. They use the Hare system, which is complicated because it is iterative. So, there were Steve and I with all these ballots. We had to arrange them in certain ways before we could start calculating the outcomes via the Hare system.

While we were doing this, I was complaining mightily about what a dopey waste of time this was. Steve turned to me very sternly and said, "Stop joking around. This is important—it's faculty politics." I knew at that point he and I were headed in different directions in our careers. Of course, the intervening 45 years have proved me completely correct.

Sam Behseta: Among all the pieces you've written in the last 25 years for *CHANCE*, do you have your own greatest hits?

Howard Wainer: I'm not sure. I'd have to look at all of them. How do you pick your favorite child? But I will say something else that is allied to this. I believe writing

this column for *CHANCE* was the single best thing I've done in my career—for my career. Because what it's done is forced me to write something intelligible every three months. As you know, I have collected 15 to 20 of these essays together every few years, reorganized them, rewritten them a little bit, written some interstitial material, and published a book. This has worked for four or five books so far. I don't think I would have done this without *CHANCE*.

Sam Behseta: Meanwhile, you've been teaching constantly.

Howard Wainer: Well, I would say teaching intermittently. I've taught for the last 10 years—one course a year at Penn. I taught for a couple years at Princeton. When I was actually in the academic business, I was at Chicago and taught there full time.

But I view teaching the same way I view cooking. I enjoy cooking one meal a week. You can plan it, shop for it, work hard at it, put your heart into it, and then watch the enjoyment other people take in your efforts. Twenty-one meals a week is not for me. I think then you're just throwing the stuff on the table.

I've taught one course a year for the last 10 years. I've worked hard at it, and I've enjoyed it immensely. The time came this year to cut back so I thanked Penn for putting up with me for all this time and stopped. They have a really wonderful department. I don't know who ranks such things, but I would put them in the highest tier.

I'm honored that they let me be a small part of the department for this past decade. But now the time has come to stop. There's a separation of three generations now between the students and me. They don't get my jokes or many of my references. The students starting there this fall were born in 1995; most of my sport jackets are older than that. I have had fully adult women in my classes who were too young to date my children. That signals it's time to stop.

Sam Behseta: Have you ever tried to use your pieces in *CHANCE* in your classroom? And what kind of feedback have you received?

Howard Wainer: It's always been positive. In fact, there are several pieces that were coauthored with students. There was a piece on exploratory data analysis I did with a student named Danielle Vasilescu that was about overcrowding among nations. Then there was a piece I did with another student, Grace Lee, about the choice of independent variables required by dating services. It began when students asked why some dating services asked to include the length of your forefinger. No one could figure out why, but some of the girls were saying there was an old wives' tale about the relationship between the length of the forefinger and the lengths of other body parts. Others suggested men often lie about their height, and finger length might provide a check on it.

Perhaps the dating service was going to use the length of the forefinger (Who's going to lie about that?) as a proxy for height. So, Grace went and gathered forefinger lengths and heights of a number of students, looked at the correlation, and calculated the standard error. It turned out that finger length isn't much of a predictor, because the standard error of the estimate of your height from your forefinger yields bounds of about plus or minus six inches, and that wasn't helpful.

Sam Behseta: One thing I tell my students is that statisticians should be able to write well because that's what they do. They analyze data and then they write about it. And it often comes as a surprise to them because their thought process is, "I'm in a quantitative field; I'm not in literature or anything like that." What's your thought on that?

Howard Wainer: It's come as a surprise to me, too. When I applied to college, I eliminated all colleges that required an essay. So, I ended up going to Rensselaer Polytechnic Institute. I hated writing because I wanted to do math. So, the joy I take in writing now astonishes me. But of course, I think one's success as an academic is related to many variables. Horsepower, of course, focus, grit, energy, good taste in problems, and the need to tell stories. You have to tell stories.

You have to construct some kind of narrative around what it is you're doing. No one's going to pay attention to your stories unless you're interesting and grammatical. And of course, the more you write, the more practice you get, and the better you are at this. When I was at Chicago, one of the advantages of being a faculty member was that you could sit in on any course you wanted. So, when I learned there was a course on writing Saul Bellow gave, I jumped at the opportunity to audit it.

One of the things I've observed about mathematicians is most of us are open to criticism of our mathematics, but not our writing. The response is usually, "Who are you to criticize my writing?" But taking this course with Bellow, the "Who are you?" is easy to answer. He has a Nobel Prize, a Pulitzer, National Book Awards, etc. So, I would write things, and he would give them back to me with comments. I would pay attention.

He always referred to writing as his craft, never as his art. He explained the reason he has had success is that he worked hard at it. Of course, he also has some game. So, I work hard at it. I write and rewrite. It's fortunate I enjoy reading my own work so much that I can reread the same thing several times and constantly pick away at it to try to find *le mot juste.* So, I emphasize the importance of clear writing to all those with whom I work.

Reading broadly is sometimes daunting. For when you read something written by a real pro, the tendency is to throw up your

Books and More Books

Since this interview, Howard Wainer has retired, and we wanted to know what he has been up to. Of course, it involves books.

Name one or two favorite blogs or books you have read.

The only blog I ever read is Andrew Gelman's, but even that only episodically. The top four books currently on my bed-stand are:

1. Ari Shavit's marvelous *My Promised Land*
2. Volume 18 in Patrick O'Brian's spectacular Aubrey/Maturin 20-volume series of continuing sea-faring novels
3. Cass Sunstein's *Impeachment*
4. Walter Isaacson's *Leonardo Da Vinci*

What have you been doing to fill your time since you retired?

Nothing has changed except I no longer commute to Philadelphia (I don't miss it) and I no longer receive a pay check (I miss that).

I have been working hard at completing the graphics history book, and I spend a fair amount of time as an expert witness (mostly in trials of teachers accused of helping their students cheat). Since there is no place in heaven for those who prosecute teachers, I always defend. School districts always seem to make the same statistical errors in conducting their investigations. I have been thinking of putting together my testimonies into some sort of guide for both teachers' unions and school districts.

I also spend a fair amount of time in the gym and, when the weather is fine, sailing on the *Surprise* (my Alerion 28).

My new book [with Michael Friendly] was tentatively titled *On the Origin of …*, but that title is almost certainly to be changed (too Darwinian). I don't know what it will be. It is planned to be published in April 2019 (if Friendly and I have the wit and energy to get it done by then). It is sort of a history of graphics—told mostly as a series of stories about the origins (and originators) of graphic forms. We begin with cave drawings of mammoths from Paleolithic times and how they influenced our understanding of what mammoths looked like and move forward through the 17th century (van Langren's one-dimensional plot of longitude) to Playfair, Minard, Snow, and many others. We also include W.E.B. DuBois and his collaboration with Minard in graphing the Great Migration. There is also a very nice chapter, principally Michael's, on the history of graphing motion that shows how we finally understood how horses run and cats land on their feet.

hands in surrender and think, "I could never do that." That may be true, so you have to focus on clarity instead. Style may come later, but clarity must come first. There are lots of wonderful statisticians who wrote terribly. How many papers/books begin with such catchy phrases as "Let I be an index set." A long way from "Call me Ishmael." John Tukey was perhaps the most influential statistician of his time, but no one would confuse his writing with Hemingway's.

Sam Behseta: Did you know him personally? How did you interact with him?

Howard Wainer: I was a student and he was a faculty member. And then I subsequently worked with him for about 30 years. When I was at ETS, I would see him frequently. He was a remarkably generous man. And so, when I had an especially thorny problem, I would go talk with him about it. He'd always have something valuable to offer.

One important lesson he taught to so many of us was about the evil of hubris. He said hubris is a bad thing because—he didn't say this, but it was certainly implicit in what he was saying—because people are—the phrase he used was—"people are different." And you don't learn anything if you don't both listen to other people and respect what they have to say. He was probably the best hubris destroyer you could imagine. You could go talk with him about anything and quickly realize you're not as smart as you thought you were. John frequently understood your problem better than you did before you arrived, but if he didn't, he certainly would before you left.

As nearly as I can tell, John knew pretty much everything. There's a wonderful story about him that I believe comes from Colin Mallows. John would sometimes get annoyed because people would ask him ridiculous questions just to see if he knew the answer. So, Colin suggested, if you want to find out how to milk an elephant, don't ask John directly. Just come in and start talking about elephants in general and eventually, when the conversation gets around to how to milk one, he'll tell you.

Sam Behseta: Were you involved in the whole discussion between Bayesianism and frequentism, and did you take sides early on, or did it not matter to you?

Howard Wainer: No. I have gradually become what is pejoratively referred to as an opportunistic Bayesian. That's if it works, I will use it, and if it doesn't, I won't. But certainly, my training was never from a Bayesian perspective. In fact, only recently did I learn why John never seemed to be a fan of Bayesian methods. This attitude was a mystery to me. I couldn't understand why Tukey wasn't more positive toward Bayesian methods as they became more practical.

Recently, the book *The Theory That Wouldn't Die,* about Bayes' Theorem, explained that Tukey was using Bayesian methods during the Cold War period on matters of national security, but he chose not to make it public because it conferred a tactical advantage to the United States. I guess that carried over into the predictions of election outcomes for NBC as well.

But I think he probably, and this is just a surmise on my part—I don't pretend to understand him—but I have a feeling he was against dogmatism because it stood in the way of the flexibility required for a first-rate scientist. And many of the Bayesians of the period were certainly dogmatic. So, he might have been in favor of Bayesian methods, but

not Bayesians. If you remember statistics in the '50s, it really was almost warfare. My friend Sam Savage tells how his 9-year-old brother was accosted at some Stanford faculty party by somebody who told him, "Your father is deeply deluded." He was talking to a 9-year-old and yet he felt it was important to point out the moral weaknesses of a Bayesian attitude.

Sam Behseta: I guess you had to identify yourself with a camp back then. And it didn't make much sense because I think, at some juncture, it was an ideological warfare between the two camps.

Howard Wainer: Well, it's almost over now. Because there are problems you can't solve any other way.

Sam Behseta: At this juncture, with the whole big data and the immediate interest in solving problems, which are at the core of science, it is possibly more efficient to be a pragmatist.

Howard Wainer: Yes. The Shibboleth is, "Does it work?" Fifteen years ago or so, I teamed up with a very young Eric Bradlow (who has since become an eminent professor at Penn). He was a dyed-in-the-wool Bayesian, and he taught me a great deal. We did a book together that was a fully Bayesian approach to doing test theory, and it forced me to learn (and eventually love) the stuff. At the time, I wasn't fully conversant with the estimation methods that made Bayesian procedures practical. MCMC [Markov chain Monte Carlo] became my favorite initials.

Sam Behseta: When I talk to PhD students, sometimes they think every PhD in statistics, when they finish their education, they ought to go into academia. But my story is, listen, actually that's a very small minority. A good chunk of PhD graduates will end up in other sectors and can be extraordinarily effective in shaping policies. How rewarding has the experience of working outside academia been to you?

Howard Wainer: Well, I'll tell you. I certainly came out of graduate school with the same point of view that you expressed—I was training to become a professor. For no good reason, other [than] who is doing the training—professors. It also seemed like a nice life. So that's the direction I wanted to go.

I found, as I pursued this path, there were aspects of being an academic that are very attractive—certainly the teaching part. The part I didn't like, because it was too difficult—at least for me—was discerning what were good problems to work on. When I first got out of graduate school, I went to Temple University and I was unhappy. I thought, well, that's because Temple isn't Princeton. I left quickly and got a job at The University of Chicago. Chicago is as good a university as you can find. Chicago treats its faculty as well as you can expect to be treated. It's a wonderful place, and the colleagues I had were terrific. And still I wasn't happy.

I found that there was an enormous amount of intellectual horsepower being used to solve trivial problems. That wasn't everybody, obviously. There is certainly wonderful and important work that comes out of academics, but since I had yet to develop good taste in the choice of problems, the university life was ill suited to me—at least at that time. So, I left Chicago and went to Washington. I worked there during the Carter administration. Toward the end of my time there, I met with Don Rubin, who was—at that point—preparing to go to work for the EPA [Environmental Protection Agency] and was assembling an all-star cast to go with him. He had already recruited Paul Rosenbaum and Rod Little, and it was flattering to be considered part of such a group. And Don had some interesting ideas. But the more we spoke about it, the clearer it became that what I was interested in doing was essentially substituting (never replacing) for Don at the Educational Testing Service (ETS), which he was leaving to go to EPA. And so, I went in that direction. I even rented Don's Princeton house from him while he lived in Washington.

At ETS, we had real problems—real in the sense that if you solved one, it could have important and immediate consequences. For example, there was one problem having to do with improving the efficiency of the shipping of tests to the hundreds of testing centers. You can't send them too early or send too many extras because of security issues. You can't send them too late because, if they arrive late, that doesn't help. How do you know how many to send? At the time, they had ad hoc rules in which, on the last day that would allow normal shipping, they would send the number of tests equal to 10 percent more than the number of people who had so far registered for that particular site. If more were needed, they would overnight them at the last minute. With hundreds of testing centers, this was expensive.

Paul Holland was looking at this problem and said if X is the number who registered by shipping time last year and Y is the number that showed up eventually, we need a rule to predict Y from X. Statisticians know how to do this. It turned out that regression equation saved ETS some number of hundreds of thousands of dollars every year. That's a nice reward for doing a regression.

ETS, in their generosity, then gave the statistics group an annual grant of a fairly large number of thousands of dollars a year to support whatever research we felt was worthwhile within our group. You can see how such work could be rewarding—and we actually did other things of broader interest. There were all sorts of important and interesting problems that appeared because ETS was a data-rich environment with a lot of resources. So, I found it to be an advantage, at least for me. ■

This interview originally appeared in *CHANCE,* Volume 28, Issue 1, Winter 2015. Read the entire free article at *http://chance.amstat.org/2015/02/wainer.*

Sharon Lohr Shares Unique Path to Writing About Crime

Debbie Sniderman, VI Ventures LLC

"Even though my path has been somewhat random, I'm happy it led here."

When she was young, Sharon Lohr's father would wake her up to watch the *Apollo* launches. He sold hydraulic equipment and was proud that one of his systems ended up being used in the Gemini program.

She followed her father's example, reading everything that came her way. She especially loved science fiction and Robert Heinlein books, in which the admirable characters shared a love for and determination to excel in mathematics.

Lohr says there were many challenges in her career, but also many good people along the way. Finishing college was a significant turning point in her life. Her mother thought girls didn't need to go to college, and Lohr was only allowed to attend after her father promised she would study music. So during her first year at Calvin College, Lohr studied piano, but also took electives in calculus and chemistry.

When her father suffered a disabling stroke during her freshman year, Lohr's mother wanted her to quit college. Her best friend's parents insisted she finish and helped her find a way to do so. By combining Social Security benefits, her savings from working during high school, and earnings from working full time during the academic year and multiple jobs in summers, she was able to graduate debt free while still helping out financially at home.

Entering Statistics

Lohr first encountered statistics during her senior year and credits her undergraduate adviser, Thomas Jager, for steering her toward it in graduate school. "He suggested I apply to statistics programs because I had done well in real analysis and wanted to learn about lots of different things," she says. "Even though I didn't really know what statisticians did, it sounded interesting, so I applied to graduate statistics programs."

Jager also recommended Lohr for a summer internship at NASA Goddard Space Flight Center. For someone who grew up reading Heinlein and following the space program, working at NASA was a dream come true.

At NASA, Lohr worked on a project to improve the spatial resolution of numerical weather predictions. Calculations for a forecast often completed after the weather occurred due to the the computer power of the time. Lohr modified front-end computer code so calculations could be sped up by the new CYBER 205 high-speed vector processor. "Figuring out how to vectorize everything and seeing applications of math, science, and statistics was interesting and fun," she says.

Lohr pursued her PhD at the University of Wisconsin-Madison and, in her first course—Statistical Experimental Design for Engineers—encountered William G. Hunter, who became one of her many mentors. She enjoyed his practical experience solving problems, designing experiments, and working with clients She also got from him a picture of what a statistician should be. At Madison, she also had the opportunity to meet W. Edwards Deming.

Launching a Career from a Workshop

Lohr began working on survey sampling and crime statistics through a chance opportunity. Her dissertation was in sequential

Lohr was the first (and still only) female Deming Lecturer. She is a Fellow of the American Statistical Association, an elected member of the International Statistical Institute, and the inaugural recipient of the Gertrude M. Cox Statistics Award for contributions to the practice of statistics.

analysis, but as a new faculty member at the University of Minnesota, she was assigned to teach the sampling class. When a colleague told her about a summer workshop on the National Crime Victimization Survey sponsored by the Bureau of Justice Statistics, she jumped at the chance to learn more about sampling to help make the subject come alive for her students.

The workshop ended up changing her career path. Seeing the problems BJS and the Census Bureau faced in conducting a survey about crime, the ingenious solutions they came up with to solve them, and the steps they took to continually improve their surveys planted in Lohr the desire to work on these problems herself.

Academic Research

Lohr joined Arizona State University as part of a two-career couple. While teaching classes and working toward tenure, she wrote numerous research articles and turned her sampling notes into her first book: *Sampling: Design and Analysis*. It was at this time that advice from Lynne Billard and colleagues in the first *Pathways to the Future* workshop helped her tremendously.

Lohr began a long-term collaboration with Jon Rao in 1996. She calls him one of the truly great statisticians of our time, with a gift for anticipating what the important research problems will be 20 years from now and solving them in the present.

After 25 years of working in academia, Lohr visited Westat to work on a crime survey that could be conducted by mail. She was subsequently invited to join the company as vice president. She enjoyed developing survey designs and statistical analysis methods for applications in transportation, public health, crime measurement, and education over the next five years.

LEFT PHOTO
Sharon Lohr receives the 2014 Deming Lecturer Award with 2015 ASA President David Morganstein.

RIGHT PHOTO
Sharon Lohr joins Jon Rao on a tour during the 2009 Conference on Small Area Estimation in Elche, Spain.

Writing About Crime Statistics

Having mostly written technical articles for statistics journals, Lohr finds writing for popular audiences to be a new challenge. Her forthcoming book, *Measuring Crime: Behind the Statistics,* is a guide for nonstatisticians on how statistical reasoning applies to crime statistics.

Lohr tells where crime data come from, how the data become statistics, and how to judge the quality of a crime statistic in a newspaper. She also tells stories about how statisticians have solved problems of measuring aspects of crime and improving the statistics.

"It's written as a general book for anyone interested in crime or statistics: students, policymakers, journalists, and interested citizens," Lohr says. "As part of the ASA series on Statistical Reasoning for Science and Society, it will emphasize how central statistical thinking is to basically everything we do."

Advice for Others Interested in Getting into Statistics

To young statisticians, Lohr advises to not plan life rigidly. "Doing so can lead to disappointment when things don't work out exactly as planned," she says. "Instead, be open to opportunities that come along. You may not think you want to work in a new area at first, but learn about it and give it a try. It may turn out to be your true calling."

Her next piece of advice is to learn about designing experiments and surveys. "Some areas of statistics are devoted to fixing things that go wrong," she says. "Persons in design can set up the study so researchers and policymakers get the information they require without needing heroic data analyses."

She also advises engaging in activities other than statistics once in a while. For Lohr, it's music. She's the only person she knows who paid for her math degree in part by playing piano. "Being a vocal accompanist was a good background for becoming a statistician, since you need to have the same sort of ability to collaborate and to anticipate what other team members will do or need," Lohr says. "It also gave me a lifelong love of opera."

"Life is short. Take advantage of opportunities to learn new things and grow," says Lohr. She can be reached at *sharonlohr.com.* ■

Megan Price: Life-Long 'Math Nerd' Finds Career in Social Justice

An interview with the executive director of the Human Rights Data Analysis Group

David Corliss, Peace-Work

Megan Price

Nonprofit and nonpartisan, the Human Rights Data Analysis Group (HRDAG, *www.hrdag.org*) is engaged in the analysis of human rights violations around the world. As the executive director of HRDAG, Megan Price designs and manages statistical projects in many locations. She is a lead investigator for a number of HRDAG projects in addition to her roles in strategic direction, management, and study design.

Tell me about your early background. Were you always interested in statistics?

I was always a math nerd. My mother has a polaroid of me in the fourth grade with my science fair project … It was the history of mathematics. In college, I was a math major for a year and then switched to statistics.

I always wanted to work in social justice. I was raised by hippies, went to protests when I was young. I always felt I had an obligation to make the world a little bit better. One event I remember is when I was in the fourth grade and there was a teacher strike where I lived in West Virginia. My parents talked about it and the issues with me at home. It was first time I remember being part of it, being taught about the grey areas, looking at the hard questions and trying to find answers.

What did you study in college? Did it do a good job of preparing you for this work?

My bachelor's and master's in statistics are from Case. While there, I became involved with the department's statistical consulting program, working at the Cleveland Clinic there in public health. I was always careful to choose elective courses in public health and social justice while studying statistics.

At Emory, where I earned my PhD in biostatistics, there is a certificate program in human rights. I took elective courses in health and human rights and an introduction to public health law class.

You once said you were "lucky to get your dream job right out of grad school." What were the important turning points you remember that had the biggest impact on your career?

There were three. I first heard about HRDAG while working on my PhD at Emory. When I was just starting to think about where to go after my doctorate, HRDAG received a grant and posted a position for a statistician who would be working on a complex sampling problem and studying the historic archive of the National Police in Guatemala. I felt like I wasn't ready yet because I wasn't certain when I would finish my dissertation, successfully defend it, etc. I had a general feeling of unreadiness for a full-time job. David Banks, who was a mentor for me, really encouraged me to apply. That was the first big turning point, when I started working for HRDAG.

The second was when HRDAG went independent. Patrick Ball started HRDAG as a AAAS project. Later, we were part of Benetech for nine years. In 2013, HRDAG spun off to become more of an independent organization and I became director of research. It was a time of change in the organization, and I took on a leadership role. Patrick was an important mentor for me at this time, helping me develop leadership and organizational skills.

The third big turning point was in 2015. Patrick and I took a look at the projects and tasks to be done, the roles to be filled, our skill sets, and where we saw HRDAG going as we moved forward. We decided to basically switch jobs: I became executive director and he became director of research.

Looking back on your career, what advice do you have for someone in school or just starting out?

When in school, be sure to take electives that are interesting. Try different things and take advantage of different options and opportunities to see what is best for you.

Go to as many conferences and workshops as possible. I know resources can be limited, so get support from wherever you can. Talk to your department at school. Many conferences offer scholarships, so apply for them. Also, your local ASA chapter can be an important source of support. Develop a strong connection with your local chapter. They can help with mentoring, resources, networking, and support for attending conferences and seminars. Be active in the ASA. That's how I made connections with many people who have helped and encouraged me.

It isn't something I'm naturally good at, but it is something important I have learned: Use Twitter and hashtags. Now I use hashtags all the time—#Data4Good, #AI, #AIRoleModels … Twitter generally and specific hashtags can be a way to learn about new things.

For people early in their career, be sure to try different things. Make sure you keep your net really, really wide open and just see what pops up. ■

Megan Price 411

- Executive Director, Human Rights Data Analysis Group, 2015 – Present
- HRDAG Director of Research, 2013–2015; Statistician, 2009–2013
- PhD, Biostatistics, and certificate in Human Rights, Emory University
- BS and MS, Statistics, Case Western Reserve University

You can follow Megan on Twitter @StatMegan

BASS XXV Scheduled for Fall

The 25th Biopharmaceutical Applied Statistics Symposium (BASS XXV) will be held October 15–19 at the Hotel Indigo Savannah Historic District in Savannah, Georgia. One-hour tutorials on diverse topics pertinent to the research, clinical development, and regulation of pharmaceuticals will be presented by speakers from academia, the pharmaceutical industry, and the US Food and Drug Administration. Two parallel, two-day short courses will be presented October 17–19. BASS will also offer a poster session.

BASS is a nonprofit entity established to support graduate studies in biostatistics. For further information, contact the BASS registrar at *rewhitworth@gmail.com* or BASS chair, Tony Segreti, at *segretia@bellsouth.net*. Visit the BASS website at *www.bassconference.org*.

TECH LEADERS
Discuss Working in Industry

Statistician and data scientist continue to be ranked among the top jobs, most recently by *U.S. News & World Report* in 2017, which listed statistician as #1 on both its lists of Best Business Jobs and Best STEM Jobs, and Glassdoor in 2018, which had six analytics and data science jobs included in its 50 best jobs in America. With this in mind, we asked leaders in industry to answer the following set of questions to help students and statistics departments better prepare for jobs in the technology industry.

~ Steve Pierson, ASA Director of Science Policy

Olivia Liao leads the pricing team at FLYR to build the next generation of revenue management systems for the airline industry. Previously, she led the data science of Demand Intelligence at Uber, which models and predicts how riders choose different Uber products. She earned her PhD in statistics from Stanford University.

Jeremy Gu is vice president of the San Francisco Bay Area Chapter. At Uber, he is a data scientist on the Experimentation Platform and works on the production of experiments. Previously, Gu was an applied scientist at Amazon, working on automated advertising and Amazon Web Services teams. He earned an MS in statistics from the University of Washington and a BS in mathematics and statistics from the University of Minnesota.

Demand Intelligence and Experimentation Platform

Please describe your career experience.

Jeremy: I chose statistics as my major because I loved analyzing data and making predictions. In high school, I always wondered how the big lottery worked. At the University of Minnesota, my first statistics class was STAT 101, taught by Sanford Weisberg and his teaching assistant Christina Knudson. They were very good at putting complex concepts into easily understood stories. So, I found their class fascinating and then was convinced statistician would be a great career for me.

After consulting many classmates, teachers, and people from industry, I decided to pursue a mathematics major, as well. When taking upper-level statistics classes, I realized mathematics, such as advanced linear algebra and real analysis, was critical for me to better understand statistical theories. Further, I took C++ and algorithms classes, because computing was a growing component in statistics curriculum.

For a college student, GPA is important. Many of my friends went to graduate school and the admissions committee valued high undergrad GPA. However, a more important component of building my career vision was talking with people from various backgrounds. I was lucky to meet Lynn Argetsinger

at the University of Minnesota when she organized events for the college of liberal arts. I attended the class reunion for the statistics school in 2011 and met many remarkable people who also graduated with a statistics degree. The advice from Stephen Eick, Denise Harbert, and Lynn Lin still influences my thinking and approach.

One of the most valuable pieces of advice I received was to look for a summer internship. I was notified by Gilad Lerman in the mathematics department with an offer to intern at a top research center working with Jiaqi Yang at Schlumberger-Doll Research in Boston. I still remember Lerman worked many hours to help me in the application process between the university and company, and I am still extremely grateful for his help. After securing that first internship, many things in my life got much easier.

I intended to pursue a PhD in statistics and, for that reason, I went to a top statistics master's program in Seattle designed for later PhD applications. At the University of Washington, the size of the master's program was tiny, and I cherished the time with my six classmates in the two-year studies. In addition to the problems discussed in the textbook, I had the opportunity to work with real data sets at school. Later, Caren Marzban offered me a research assistant position and we published a paper together.

Tech giants Amazon and Microsoft were growing rapidly in the Seattle area, and I had opportunities to meet people at both companies. As a result, my view of my career changed and I decided to explore job opportunities in the Seattle area. At that time, big data and machine learning were two big directions, and many people started to use data science for jobs working with numbers and models. After getting my MS in statistics from the University of Washington, Amazon hired me in the cloud computing department. One year later, I switched to another team at Amazon focusing on machine learning models on the automated advertising. Now I work on the experimentation platform team at Uber and am still a student who needs to learn and grow.

What do you like about working in the technology industry? What are the challenges?

Jeremy: As a data scientist, I enjoy working in the technology industry because there are many opportunities to apply science techniques to solve challenges. At Amazon, I first worked in the cloud computing business and then learned about many things, including Spark and Hadoop. On the business side, I honed my skills in advertising and marketing. For example, a data scientist can design an online experiment to help choose the right metric to measure an advertising campaign. Moreover, data scientists must know how to build machine learning predictive models in an end-to-end fashion; the entire process includes getting the data, training the model, putting the model into online production, and measuring and improving the performance. Outside the data science domain, you will learn many other interesting problems and meet many talented co-workers from other backgrounds.

As with any other job, there are also challenges with working in the tech industry. First, the tech industry changes very fast and data scientists need to keep up with the latest technologies. For example, after graduating in 2014, I spent another year studying Hadoop Ecosystems and Spark in a certificate program. I found it useful because I didn't use Hive, Hadoop, or many other new tools on big data platforms at school. Coursera was a great resource for me as a student. I took both free and paid classes in machine learning, big data, digital marketing, and so on. I also recommend attending local meetups to connect with people with similar interests. Sometimes, studying with a study buddy motivates both people.

There is a big difference between learning what's in textbooks and being able to apply those concepts to industry work. It can take several years for new graduates to adapt to the professional environment. I refer to this time as post-education education. I learned as much during this time as I did at school. For example, I learned how to use several SQL languages on different platforms. Data querying, parameter tuning, outlier detection, and many other techniques have quick and clever solutions once you acquire enough experience. Besides technical knowledge, students entering the workforce will learn how to work with a direct manager, communicate with people from various backgrounds, manage personal income, and many other things that aren't taught in a classroom.

How is the demand for statisticians in the technology industry? What are the main degrees you consider when looking for candidates with statistical expertise?

Olivia: The demand for statistician remains high, sometimes even when technology companies don't realize it. In the era of big data, the industry needs statisticians with keen eyes to find insightful business trends from messy, often confounded data, as

well as to determine how the company's products affect behaviors through causal inference methods. Statisticians with skills in high-dimensional predictive models, which sometimes are the core products of the company or team, are also in high demand. Also, even with big data, the business-relevant events data scientists deal with can be rare (e.g., ads click-through, client default on loans, insurance claims, etc.) or the data become sparse when you dig into particular business segments. The ability to handle these analytical challenges is highly desired.

Most hiring managers consider candidates with degrees in statistics, biostatistics, economics, and industrial engineering. However, the exact degree is less important than relevant internships and projects. Also, different companies and teams require a diverse background for data scientists. While A/B experiment and regression modeling are common skills required for all positions, not all teams need experts in machine learning or deep learning. Many look for generalists who have broader skill sets to take on any challenging analytical problems, while some look for experts in causal inference, time-series forecasting, spatial-temporal modeling, etc.

What do you see as the most important statistical skills in the technology industry? What are other important skills necessary for a successful career in this sector?

Olivia: The most important statistical skills we practice daily and are looking for in candidates include experimentation (hypothesis testing), data visualization, regression, and high-dimensional modeling. The first two are often overlooked. However, in our line of work, we are often given vague business problems and need to lay out plausible hypotheses (e.g., factors that drive consumer purchasing decisions). Then we devise experiments—either online, simulated, or observational—to validate those hypotheses. The results can sometimes fundamentally change the direction of the product development and roadmap of an entire team.

To establish a successful career in this sector, data scientists also need strong communication skills. Our work frequently requires cross-functional collaboration with multiple groups. The value of data science will not be fully realized until others understand the value of the work, the complexity in the algorithm, and the exact problem we are trying to solve. "Singular value decomposition" does not convey the problem, and "increase AUC" is not sufficient to get people aligned on the value data scientists bring to the table. Moreover, being able to clearly articulate to your manager any difficulties you are facing can help them resolve your concern most efficiently.

In the industry, data scientists are often asked to tackle vague business problems, rather than textbook-style exercises, and the ability to break down a large problem into smaller pieces is essential. The first step is identifying the underlying business action. For instance, if your project is designing a restaurant rating system, the solutions will differ based on whether the rating is being used merely as a reference for consumers, to rank restaurants in the recommendation system, or to determine prices. From there, a series of sub-questions to address include the type of data needed for the analysis or modeling, any assumptions needed to simplify the problem, potential low-hanging fruit approaches, the success criteria, and how to measure them. To successfully carry out the project, you would also need to identify all the tasks and/or blockers and the best person (not just yourself) to work on each. Together with the team, you decide whether the tasks can be carried out in parallel or need to be sequenced.

Finally, almost all data scientists and statisticians in the tech industry need to code. The task may be as simple as developing an automated metrics monitoring dashboard or as challenging as implementing the model in the production system. In smaller companies, you may even be involved in guiding the design of data warehouse infrastructure.

What advice do you have for students interested in working in the technology industry? Any advice for students in general?

Jeremy: I have two pieces of advice for students looking for opportunities at tech companies.

The first piece is a general one and something many teachers and parents might say: Do your best to study. A solid understanding on statistics, machine learning, and/or big data will be rewarding when working at tech companies.

The second piece of advice is to ask yourself what your strength is and what your interest is. When the two items align, your career path will be rewarding no matter if it is in data science or not. In my view, data science is not a fashion, but a rigorous science after years of hard work from mathematicians, computer scientists, statisticians, economists, and many others. Those people made huge progress in their fields in the past and are still working hard to advance in the new AI/ML technology. I suggest spending a large amount of time thinking about your goals before committing to a data science career.

What advice do you have for statistics and biostatistics departments?

Jeremy: I have received many messages from statistics majors on LinkedIn asking for job advice. I put the messages into the following three buckets:

1. Questions about how to choose classes toward a data science career
2. Questions about how to apply for data scientist positions
3. Questions about how to prepare for job interviews

I believe many of those students didn't collect enough satisfying answers from the school and therefore asked me instead. I would make the following proposals to the program adviser at a university:

- Encourage students to attend networking events and conferences
- Partner with tech companies to provide internship opportunities to students
- Offer course guides on data science to students looking for a data science job
- Invite people from the industry to give frequent seminars or guest lectures to help students understand the industry needs and opportunities

What opportunities for advancement and professional growth exist for data scientists and statisticians in industry, and what advice for young professionals would you have to take advantage of those opportunities?

Olivia: Career growth in the industry can generally be divided into two tracks: individual contributor and manager. This separation is more apparent in large technology companies. Larger companies also have more clearly defined career trajectories and resources for self-development. However, smaller companies like startups usually provide more and faster growth opportunities.

People typically choose the individual contributor track if they are passionate about driving business impact through deep technical expertise and profound domain knowledge but don't want to spend as much time on people management. In comparison, people with a desire to influence the company through formal leadership roles or team building are ideal candidates for the manager track.

To grab these opportunities, we encourage young professionals to start thinking about their future growth early on in their careers and establish a habit to periodically (e.g., every six months) evaluate their progress. Opportunities don't just fall into the lap of the hardest-working individual. Proactively discuss opportunities with your manager or mentor and ways you can achieve them. We also encourage young professionals to think beyond technical skills development and constantly practice soft skills—including communication, leadership, and team coordination—to effectively move projects forward. No matter which track you pick, the ability to help others achieve collective goals and deliver the greatest value is critical to a successful career.

Risk Management

Xin Ge is chief risk and analytics officer for Afterpay Touch Group, responsible for risk, data, and analytics initiatives globally. Before Afterpay, Ge was head of risk management at Uber and VP of global seller risk management, resolutions, and protections at PayPal. He was also the founder of PayPal Risk Management Center in China. Ge started his career on risk management at eBay 15 years ago. He holds a BS in computer science from the University of Science and Technology of China and PhD in statistics from The University of North Carolina at Chapel Hill.

Please describe your career experience.

My first job out of school was as a consultant at Insightful Corp. in Seattle. We used S-PLUS to simulate high-frequency trading algorithms for various trading firms. The work was fun and matched well with my PhD thesis, which was related to derivative pricing. It required a good understanding of financial products, as well as strong communication skills to interact with the clients. To create the most efficient equity portfolio, we often needed to solve optimization problems.

I joined eBay in 2002 and spent the next 14 years taking various risk management roles with eBay and PayPal. I started as a statistician at eBay, working on metrics, reporting, and modeling to tackle fraud and abuse in the eBay marketplace, including account takeovers, merchant fraud, unpaid items, and shill bidding. I took on management responsibilities gradually, starting with technical functions at eBay and growing into more general management where I was responsible for end-to-end business results at PayPal. From 2010–2015, I spent five years in Shanghai and built PayPal's Risk Management Center in China. At Uber, I led

the global risk management team, working with a cross-functional group of product managers, engineers, and data scientists to tackle unique risk challenges related to payment, account security, and marketplaces.

Two months ago, I joined Afterpay Touch Group as chief risk and analytics officer. Afterpay, which recently joined SP/ASX 200, has had tremendous success in Australia and was named Australia's Fintech company of the year. I am thrilled to be part of the team, marching on a journey as we expand globally, beginning in the US.

What do you like about working in the technology industry? What are the challenges?

Risk management is a fascinating world, with challenging and engaging work. First, you feel great about what you do as you stand for justice, protecting the company and the needs of good guys. The work is also intellectually stimulating, because you are working against fraudsters—real people who use advanced techniques to cheat. There is a lot of back and forth since every move you make is quickly countered. You also don't have a choice but to keep up with the latest technology, because your opponents do. It has been a fun journey shifting battlefields from offline to online and mobile and witnessing the evolution of using intelligence such as IP, device, and GPS signals.

How is the demand for statisticians in the technology industry? What are the main degrees you consider when looking for candidates with statistical expertise?

As the technology world has evolved so much in the past 15 years, we have seen an explosion of data needs and a sharp increase in demand for candidates who can make sense out of data and drive business decisions. Our team members usually have advanced quantitative or engineering degrees—statistics, economics, computer science, operations research, and engineering are some of the common ones.

What advice do you have for students interested in working in the technology industry? Any advice for students in general?

The first step of any data analysis is to prepare data. In the real world, data don't always come in the ideal form. At companies such as eBay, PayPal, and Uber, data often come in an unstructured way and at a super scale. Hence, I think the most basic and important skill for anyone who starts his/her career is the ability to handle complex and large-scale data. Programming languages such as SQL, Python, and R will come in handy for anyone interested in data-related work.

Another important skill is to discover patterns in the data, whether you prefer the term predictive modeling or machine learning. The line has blurred quite a bit between traditional statisticians and computer scientists, and we have a lot more computer scientists in this field than 10 years ago. The following is a generalization and certainly not necessarily true for everyone, but I have observed that statisticians tend to have a deeper connection to the business problem, trying to link the patterns with common sense so they can interpret the results. Computer scientists, on the other hand, are usually stronger in manipulating and studying large-scale, complex data and implementing models in a production environment faster. Both computer science and statistics are important, and a combination of both skills would make someone valuable.

What advice do you have for statistics and biostatistics departments?

To prepare for industry work, it would be beneficial to add coursework and/or experience for students to process real-world data and solve real business problems. I would also encourage students to look for summer internships, because they are a valuable experience that will not only help students develop technical skills, but also identify how they can fit into the actual work environment. ■

2017 Photo courtesy of Megan Griffiths

From left: Ellen Kulinsky of the University of California at Berkeley, Sastry Pantula of Oregon State University, Aaron Huang of the University of Washington, Betsy Hensel of the University of Virginia, and Shelby Taylor of Brigham Young University. To read more about the OSU REU program, visit OSU's website (*https://stat.oregonstate.edu/reu*).

ASA REU EXPERIENCES OFFER STUDENTS VALUE

Mark Daniel Ward

In the April 2016 issue of *Amstat News*, we announced that the ASA received a three-year National Science Foundation (NSF) grant to establish a series of Research Experiences for Undergraduates (REUs). Three such REU experiences have been offered during each of the previous three summers, with four students per site.

With this issue's theme of "training the next generation of data scientists" in mind, we encourage faculty members to apply for their own REU site at *https://bit.ly/2Au9LSu*.

Only a few sites in the Mathematical Sciences are dedicated to statistics. Similarly, encourage your students to apply for REU opportunities; the full list is posted at *https://bit.ly/1xZqbYN*.

Students at the ASA REU sites are participating in research that run the gamut, from theoretical to applied to data-driven topics. For instance, Julia Sharp and her colleagues at Colorado State University designed projects for the students that included the exploration of anticancer agent properties, classification of salmonella serotype proteins, and simulating and examining optical density versus time data from particles under varying experimental conditions. At Morgan State University, Monica Jackson and Leon Woodson built on the years of success of the SPIRAL team to host the first Summer Program Advancing Techniques in Applied Learning of Statistics (SPATIALStats). Their goals included developing and analyzing spatial statistical methods that will detect geographic distributions of populations at risk, determine cancer rates, and define environmental conditions that affect diseases.

I just returned from a visit to see Sat Gupta at the University of North Carolina at Greensboro. He and his colleagues and

These students spoke about probability and statistics as if they were already in graduate school.

ASA REU Experiences

2016

Summer Program in Research and Learning (SPIRAL) at **Morgan State University**, led by Monica Jackson

Data Science for the Public Good (DSPG) at **Virgina Tech**, led by Stephanie Shipp

REU at **Lamar University**, led by Kumer Das

2017

Oregon State University, led by Thomas Sharpton

Emory University, led by Lance Waller

WINona STATe StatisticsREU (WINSTATS-REU), led by Silas Bergen

2018

Colorado State University, led by Julia Sharp

The University of North Carolina at Greensboro, led by Sat Gupta

Summer Program Advancing Techniques in Applied Learning of Statistics (SPATIAL-Stats) at **Morgan State University**, led by Leon Woodson and Monica Jackson

students focused on developing and comparing predictive models for various diseases using machine learning and robust regression techniques. They also used publicly available epidemiology data for building predictive models. Gupta worked with students on methods for ensuring survey respondents have confidentiality in their responses using randomized response models.

When I was talking with the students over lunch at UNC-Greensboro, I was particularly impressed to hear about how deeply they dived into their research projects. Throughout our meal, the students discussed the pros and cons of randomized response models with such confidence that one could easily have thought the models were all originally developed by them.

I was also struck by the degree to which the students appreciated the underlying statistical theory. For example, Purdue student Amber Young opted to work on two research projects over the summer. Amber was my student in a probability theory course a couple years ago, and I remember she enjoyed probability theory. Nonetheless, as we sat at lunch and discussed her summer experience, I was impressed with how comfortable she and the other students were about certain probability topics we usually don't cover until the PhD-level course. I brought up a few qualifying exam questions with the students, just to see how they responded. These students spoke about probability and statistics as if they were already in graduate school.

In many cases, students travel to other sites during their REU experiences (e.g., field trips to federal research laboratories, research conferences, companies (such as SAS), and other institutions) as part of their professional development experience. The REU sites are intended to foster the full career development of the student. With this in mind, I asked some of the students to reflect on their experiences over the years. Speaking about her participation from summer 2017, Zavia Epps said, "I had a really great personal and professional experience during the REU I participated in last summer at Emory University. I had the amazing chance to learn a lot about biostatistics and, truly, I think it is an amazing field—career and education wise." She continued, "Being with Dr. Waller and the department of biostatistics at Emory was really great, and I really liked every moment of it."

My colleagues and I believe the ASA REU experiences are valuable for students. We hope more faculty will apply to start their own sites. Kumer Das at Lamar University, as an example, had one of the pilot sites in summer 2016. Building on his experience, he applied for an NSF grant and was awarded with funding to start his own site at Lamar, beyond the ASA REU.

If I can assist colleagues in any way with their own proposal preparations for NSF REU sites, please reach out to me. REU programs in statistics have a tremendous student impact. We need more of them! Write to me at *mdw@purdue.edu.* ■

HIRING a DATA SCIENTIST

Mikhail Popov, Wikimedia Foundation

Editor's Note: This article was reprinted from the Wikimedia blog (*https://wikimediafoundation.org/2017/02/02/hiring-data-scientist*) and is licensed under CC BY 3.0 (*https://creativecommons.org/licenses/by/3.0*).

Note: This post applies to employers hiring data analysts, data scientists, statisticians, quantitative analysts, or any one of the dozen more titles used for descriptions of the job of "turning raw data into understanding, insight, and knowledge" (Wickham & Grolemund, 2016, *https://bit.ly/2MZS3so*), the only differences being the skills and disciplines emphasized.

We recently needed to backfill a data analyst position at the Wikimedia Foundation. If you've hired for this type of position in the past, you know it is no easy task—both for the candidate and the organization doing the hiring.

Based on our successful hiring process, we'd like to share what we learned and how we drew on existing resources to synthesize a better approach to interviewing and hiring a new member of our team.

Why Interviewing a Data Scientist Is Hard

It's really difficult to structure an interview for data scientist positions. In technical interviews, candidates are often asked to recite or invent algorithms on a whiteboard. In data science interviews, specifically, candidates are often asked to solve probability puzzles that seem similar to homework sets in an advanced probability theory class. This shows they can memorize formulas and figure out the analytical solution to the birthday problem in five minutes, but it doesn't necessarily indicate whether they can take raw, messy data and tidy it up, visualize it, glean meaningful insights from it, and communicate an interesting, informative story.

These puzzles, while challenging, often have nothing to do with actual data or the kinds of problems that would be encountered in an actual working environment. It can be both a frustrating experience for candidates and organizations alike—which is why we wanted to think about a better way to hire a data scientist for our team.

We also wanted our process to attract diverse candidates. As Stacy-Marie Ishmael, a John S. Knight Fellow at Stanford University and former managing editor for Mobile at BuzzFeed News, put it, "Job descriptions matter … and where they're posted matter[s] even more."

How to Write a Job Post That Attracts Good, Diverse Candidates

Defining 'Data Scientist'

The most obvious (but sometimes overlooked) issue in hiring a data scientist is figuring out what kind of skillset you're actually looking for. The term "data scientist" is not standard; different people have different opinions about what the job entails depending on their background.

Jake VanderPlas, a senior data science fellow at the University of Washington's eScience institute describes data science (*https://bit.ly/2uHqI63*) as "an interdisciplinary subject" that "comprises three distinct and overlapping areas: the skills of a statistician who knows how to model and summarize data sets (which are growing ever larger); the skills of a computer scientist who can design and use algorithms to efficiently store, process, and visualize this data; and the domain expertise—what we might think of as 'classical' training in a subject—necessary both to formulate the right questions and to put their answers in context."

That's more or less the description I personally subscribe to, and the description I'll be using for the rest of this piece.

How to Ensure You're Attracting a Diverse Group of Candidates

Now that you've defined data scientist, it's necessary to move onto the next section of your job description: what a person will actually do!

The exact phrasing of job descriptions is important because research in this area has shown women feel less inclined to respond to "male-sounding" job ads (*https://bit.ly/2NZwpoH*) and truly regard "required qualifications" as required qualifications (*https://bit.ly/1zZzB3B*). In a study of gendered wording in job posts by Danielle Gaucher et al. published in a 2011

Mikhail Popov got his start with R and statistics at California State University, Fullerton, where he did undergraduate research in the application of statistics to neuroscience. He continued working with brain data as part of the master's of statistical practice program at Carnegie Mellon University, followed by his employment at the Neuropsychology Research Program with the University of Pittsburgh Medical Center. These days, Popov is a data analyst for the Wikimedia Foundation, where his work focuses on supporting teams that improve the Wikipedia reading experience. He loves brewing coffee, cooking, baking, hiking, and sharing his knowledge with others.

Journal of Personality and Social Psychology article, they found "job advertisements for male-dominated areas employed greater masculine wording than advertisements within female-dominated areas" and "when job advertisements were constructed to include more masculine than feminine wording, participants perceived more men within these occupations and women found these jobs less appealing."

We had a job description (JD) that was previously used for hiring me, but it wasn't perfect. It included lines like "experience contributing to open source projects," which could result in preference for people who enter and stay in the open source movement because they don't experience the same levels of harassment others experience or a preference for people who have the time to contribute to open source projects (which may skew toward a certain type of person.)

We consulted Geek Feminism wiki's how-to on recruiting and retaining women in tech workplaces (*https://bit.ly/1slPnWL*) and the solutions to reducing male bias in hiring (*https://bit.ly/2zSnKkY*) when rewriting the job description so to not alienate any potential candidates. From that document, we decided to remove an explicit requirement for years of experience and called out specific skills women are socialized to be comfortable with associating with themselves, adding time management to required skills and placing greater emphasis on collaboration.

Once we finished this draft, we asked for feedback from several colleagues who we knew to be proponents of diversity and intersectionality.

A super important component of this: We did not want to place the burden of diversifying our workforce on the women or people of color in our workplace. Ashe Dryden, an inclusivity activist and expert on diversity in tech spaces, wrote, "Often the burden of fostering diversity and inclusion falls to marginalized people" and "all of this is often done without compensation. People internal to the organization are tasked with these things and expected to do them in addition to the work they're already performing" (*https://bit.ly/2zWIqZc*).

We strongly believe everyone is responsible for this, and much has been written (*https://bit.ly/2mpo1SN*) about how the work of "[diversifying a workplace] becomes a second shift, something [members of an under-represented group] have to do on top of their regular job." To remedy this, we specified colleagues to give feedback during their office hours, when/if they had time for it (so it wouldn't negatively affect their work) and only if they actually wanted to help out.

From the feedback, we rephrased some points and included encouragement for a diverse range of applicants ("Wikimedia Foundation is an equal opportunity employer, and we encourage people with a diverse range of backgrounds to apply. We also welcome remote and international applicants across all time zones."). We then felt confident publishing the job description, which our recruiters advertised on services like LinkedIn. In addition, we wanted to advertise the position where DataSci women would congregate, so I reached out to a friend at R-Ladies (a network of women using R) who was happy to let the mailing list know about this job opening.

In short, be proactive, go where people already congregate, and ensure your language in a job post is as inclusive as possible and you will likely attract a wider pool of potential candidates.

Sample Job Description

You might be asking yourself, "So what did this job description actually look like?" Here it is, with important bits bolded and two italicized notes interjected:

> The Wikimedia Foundation is looking for a pragmatic, detail-oriented data analyst to help drive informed product decisions that enable our communities to achieve our vision: a world in which every single human being can freely share in the sum of all knowledge.
>
> Data analysts at the Wikimedia Foundation are key members of the product team who are the experts within the organization on measuring what is going on and **using data to inform the decision-making process. Their analyses and insights provide a data-driven approach for product owners and managers** to envision, scope, and refine features of products and services that hundreds of millions of people use around the world.
>
> You will join the Discovery Department, where we build the anonymous path of discovery to a trusted and relevant source of knowledge. **Wikimedia Foundation is an equal opportunity employer, and we encourage people with a diverse range of backgrounds to apply. We also welcome remote and international applicants across all time zones.**
>
> **As a Data Analyst, you will:**
>
> - Work closely with product managers to build out and maintain detailed ongoing analysis of the department's products, their usage patterns, and performance.
> - Write database queries and code to analyze Wikipedia usage volume, user behavior, and performance data to identify opportunities and areas for improvement.
> - **Collaborate with the other analysts** in the department to maintain our department's dashboards, ensuring they are up to date, accurate, fair, and focused representations of the efficacy of the products.

- Support product managers through rapidly surfacing positive and adverse data trends, and complete ad hoc analysis support as needed.
- Communicate clearly and responsively your findings to a range of departmental, organizational, volunteer, and public stakeholders—to inform and educate them.

Notice the emphasis on collaboration and communication—the social aspect, rather than technical aspect, of the job.

Requirements:

- Bachelor's degree in statistics, mathematics, computer science, or other scientific fields (or equivalent experience).
- Experience in an analytical role extracting and surfacing value from quantitative data.
- Strong eye for detail and a passion for quickly delivering results for rapid action.
- Excellent written, verbal, scientific communication and time-management skills.
- Comfortable working in a highly collaborative, consensus-oriented environment.
- Proficiency with SQL and R or Python.

Pluses:

- Familiarity with Bayesian inference, MCMC, and/or machine learning.
- Experience editing Wikipedia or with online volunteers.
- Familiarity with MediaWiki or other participatory production environment.
- Experience with version control and peer code review systems.
- Understanding of free culture / free software / open source principles.
- Experience with JavaScript.

Notice how we differentiate between requirements and pluses. Other than SQL and R/Python, we don't place a lot of emphasis on technologies and specific advanced topics in statistics. We hire knowing the candidate is able to learn Hive and Hadoop and that they can learn about multilevel models and Bayesian structural time series models if a project requires it.

Benefits & Perks*:

- Fully paid medical, dental, and vision coverage for employees and their eligible families (yes, fully paid premiums!)
- The Wellness Program provides reimbursement for mind, body, and soul activities such as fitness memberships, massages, cooking classes, and much more
- The 401(k) retirement plan offers matched contributions at 4% of annual salary
- Flexible and generous time off—vacation, sick, and volunteer days
- Pre-tax savings plans for health care, child care, elder care, public transportation, and parking expenses
- For those emergency moments—long- and short-term disability, life insurance (2x salary), and an employee assistance program
- Telecommuting and flexible work schedules available
- Appropriate fuel for thinking and coding (aka, a pantry full of treats) and monthly massages to help staff relax
- Great colleagues—international staff speaking dozens of languages from around the world, fantastic intellectual discourse, mission-driven and intensely passionate people

* for benefits-eligible staff; benefits may vary by location

Take-Home Task

Many engineering and data science jobs require applicants to complete problems on a whiteboard. We decided not to do this. As Tanya Cashorali, the founder of TCB Analytics, put it, "[Whiteboard testing] adds unnecessary stress to an environment that's inherently high stress and not particularly relevant to real-world situations." Instead, we prefer to give candidates a take-home task. This approach gives candidates the opportunity to perform the necessary background research, get acquainted with the data, thoroughly explore the data, and use the tools they are most familiar with to answer questions.

After our candidates passed an initial screening, they were given 48 hours to complete a data analysis task (*https://bit.ly/2LMAaAw*), inspired by an S&D data analyst task (*https://bit.ly/2NXZwrw)* I had completed during my interview process. The tasks were designed so the candidate would have to do the following:

- Develop an understanding and intuition for the provided data set through exploratory data analysis
- Demonstrate critical thinking and creativity
- Deal with real-world data and answer actual, potentially open-ended questions
- Display knowledge of data visualization fundamentals
- Write legible, commented code

- Create a reproducible report (e.g., include all code, list all dependencies) with a summary of findings

We recommend designing a task that uses your own data and a question you've answered previously to give candidates an example of their day-to-day work in the future. If your team or organization has worked on a small-scale, data-driven project to answer a particular business question, a good starting point would be to convert that into the take-home task.

Interview Questions

Now that you have your candidates, you have to interview them. This, too, can be tricky, but we wanted to judge each candidate on their merits, so we created a matrix ahead of time that could measure their answers.

We wanted to emphasize how our prospective applicants thought about privacy and ethics. From how we handle requests for user data to our public policy on privacy, our guidelines for ethically researching Wikipedia, and our conditions for research efforts, it is clear that privacy and ethical considerations are important to the Wikimedia Foundation. We wanted to ensure final candidates could both handle the data and privacy concerns that come with this job.

When we thought about the sorts of questions we've been asked in previous interviews and the kinds of topics that were important for us, we devised the following goals:

- Assess candidate's critical thinking and research ethics
- Require candidate to interpret, not calculate/generate, results
- Learn about candidate's approach to analysis
- Gauge candidate's awareness/knowledge of important concepts in statistics and machine learning

To that end, I asked the candidates some or all of the following questions:

- What do you think are the most important qualities for a data scientist to have?
- **Data Analysis:**

  —What are your first steps when working with a data set? ("Exploratory data analysis" is too vague! Inquire about tools they prefer and approaches that have worked for them in the past.)

  —Describe a data analysis you had the most fun doing. What was the part you personally found the most exciting?

  —Describe a data analysis you found the most frustrating. What were the issues you ran into and how did you deal with them?

I used the following questions to assess the candidate's ability to identify ethics violation in a clear case of scientific misconduct because I wanted to work with someone who understood what was wrong with the case, knew why it was wrong, but also could devise a creative solution that would respect privacy. First, I asked if they'd heard about the OKCupid fiasco (*https://bit.ly/1R21bFQ*). If they hadn't, I briefly caught them up on the situation, described how answers on OKCupid work (if they didn't know), and specifically mentioned that the usernames were left in the data set.

- Please discuss the ethical problems with compiling this data set in the first place and then publicly releasing it.
- You're an independent, unaffiliated researcher. Maybe you're a researcher here at the foundation, but you worked on this project in your personal capacity outside of work. Describe the steps you might take to make the individuals in the data set less easily re-identifiable and the kinds of steps you might take before releasing the data set.
- Concepts in Statistics:

  —Statistical power, *p*-value, and effect size make up an important trio of concepts in classical statistics that relies on null hypothesis significance testing. As Andrew Gelman, a professor of statistics at Columbia University, writes, "Naïve (or calculating) researchers really do make strong claims based on *p*-values, claims that can fall apart under theoretical and empirical scrutiny" (*http://andrewgelman.com/2016/03/07/29212*). I presented the outcome of a large sample size (e.g., 10K subjects) A/B test that yielded tiny (e.g., odds ratio of 1.0008) but statistically significant (e.g., $p < 0.001$) results, and then I asked if we should deploy the change to production. Why or why not?

  —Bootstrapping is a popular and computationally intensive tool for nontraditional estimation and prediction problems that can't be solved using classical statistics. While there may be alternative nonparametric solutions to the posed problem, the bootstrap is the simplest and most obvious for the candidate to describe, and we consider it an essential tool in a data scientist's kit. I asked the candidate how we might approach an

A/B test in which we developed a new metric of success and a similarity measure that we can't use any of the traditional null hypothesis significance tests for.

—Not satisfying the assumptions in statistical models can lead the scientist to wrong conclusions by making invalid inferences. It was important for us that the candidate was aware of the assumptions in the most common statistical model (*https://bit.ly/2vpstF5*) and understood if/how the hypothetical example violated those assumptions. Furthermore, we wanted to see whether the candidate could offer a more valid alternative from—for example—time series analysis, to account for temporal correlation. "One of the things we're interested in doing is detecting trends in the usage of our APIs—interfaces we expose to the public so they can search Wikipedia. Say I've got this time series of daily API calls in the millions and I fit a simple linear regression model to it and I get a positive slope estimate of 3,000 from which I infer that use of our services is increasing by 3,000 API calls every day. Was this a correct solution to the problem? What did I do wrong? What would you do to answer the same question?"

- Concepts in Machine Learning:

  —**Model Tuning:** Many statistical and machine learning models rely on parameters (and hyperparameters) that must be specified by the user. Sometimes, software packages include default values, and sometimes those values are calculated from the data using recommended formulas—for example, for a data set with **p** features in the example below, **m** would be $\sqrt{\mathbf{p}}$. A data scientist should not always use the default values and needs to know how parameter tuning (usually via cross-validation) is used to find a custom, optimal value that results in the smallest errors but also avoids overfitting. First, I asked if they knew how a random forest works in general and how its trees are grown. If not, it was not a big deal because I'm not interested in their knowledge of a particular algorithm. I reminded them that, at every split, the algorithm picks a random subset of **m** features to decide which predictor to split on, and then I asked what **m** they'd use.

  —**Model Evaluation:** It's not enough to be able to make a predictive model of the data. Whether forecasting or classifying, the analyst needs to be able to assess whether their model is good, how good it is, and what its weaknesses are. In the example below, the classification model might look good overall (because it's really good at predicting negatives since most of the observations are negatives), but it's actually terrible at predicting positives! The model learned to maximize its overall accuracy by classifying observations "negative" most of the time. "Let's say you've trained a binary outcome classifier and got the following confusion matrix. This comes out to misclassification rate of 17%, sensitivity of 99%, specificity of 18%, prevalence of 80%, positive predictive value of 83%. Pretend I'm a not-so-technical executive and I don't know what any of these numbers mean. Is your model good at predicting? What are its pitfalls, if any?"

	Predicted Positive	Predicted Negative
Actual Positive	2K	9K
Actual Negative	500	45K

It Worked!

Based on this process, we successfully hired Chelsy Xie (*https://meta.wikimedia.org/wiki/User:CXie_(WMF)*)—who writes awesome reports, makes fantastic additions to Discovery's dashboards (like sparklines and full geographic breakdowns), and (most importantly) is super inquisitive and welcomes a challenge (core traits of a great data scientist).

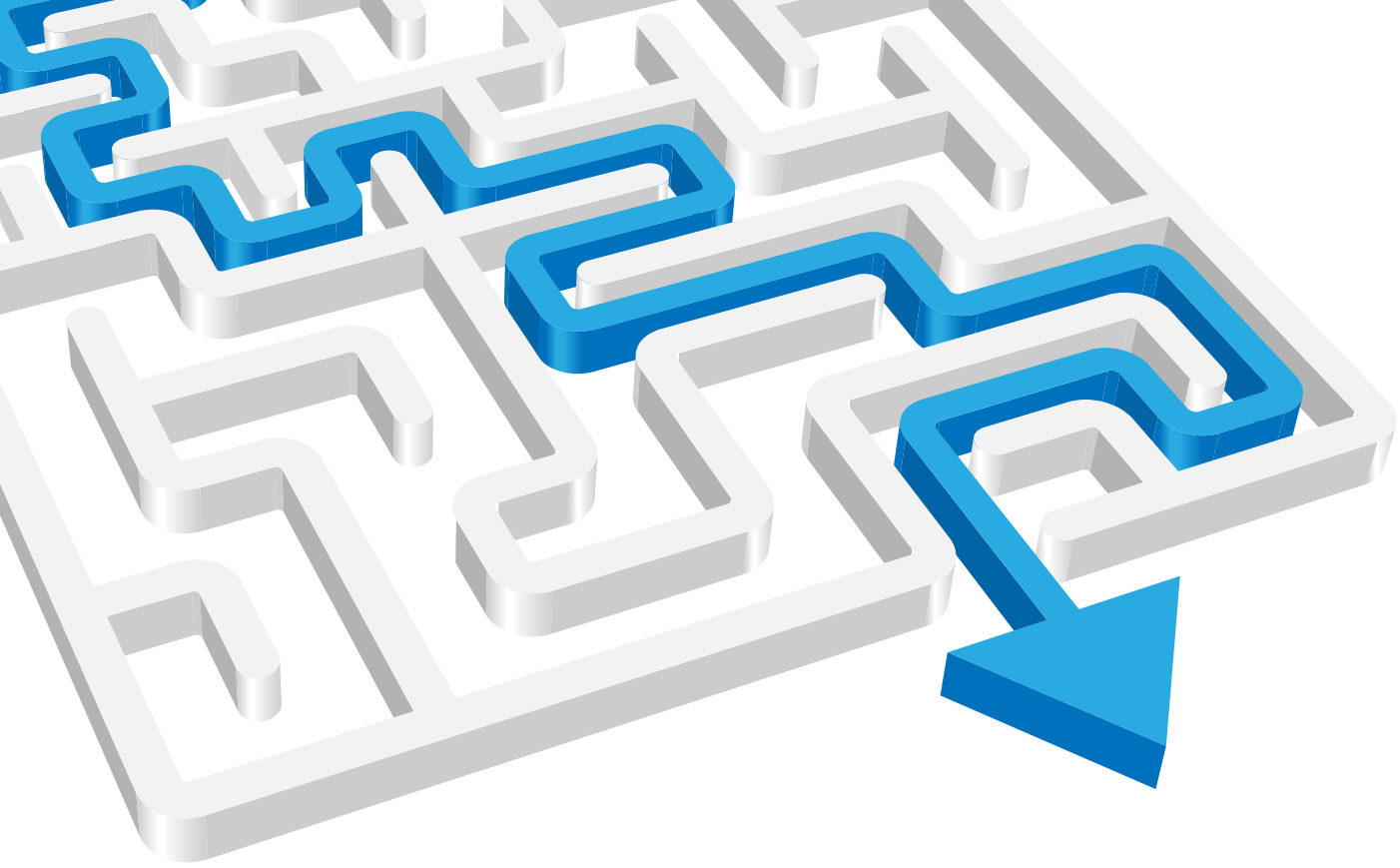
This process was easier, in part, because Chelsy was not the first data scientist hired by the Wikimedia Foundation; our process was informed by having gone through a previous hiring cycle, and we were able to improve during this iteration.

It's harder for employers who are hiring a data scientist for the first time because they may not have someone on their team who can put together a data scientist–oriented interview process and design an informative analysis task. Feel free to use this guide as a way to navigate the process for the first time, or for improving your existing process.

This isn't the only way to interview a candidate for a data scientist position, nor is it the best way. Much of our thinking on how to approach this task was shaped by our own frustrations as applicants, as well as our experience of what data scientists actually do in the workforce. These insights likely also apply to hiring pipelines in other technical disciplines.

We are also interested in continually improving and iterating this process. If you have additional tips or would like to share best practices from your own data scientist hiring practices, please share them.

References and further reading can be found on the original blog post at *https://bit.ly/2vwYefD*. ■



The Short and Winding Road Through a Career in Industry

Lyndsay Nobel has both a BS and an MS in mathematics with a focus in statistics from McNeese State University and an ABD in applied statistics and research methods from the University of Northern Colorado. She has been working as an analytics professional since 2001.

In 1995, I was a second-year, undergraduate, transfer student in the math department and needed to fulfill a state requirement for an introductory math class. Since I had already taken a calculus class, I could not take college algebra, and the only class that met the state requirement was elementary statistics (strange rules I still don't understand, but what a fortuitous happenstance). I was drawn into the subject matter and haven't looked back.

When I reached the end of my bachelor's program, the only job opportunities I was qualified for were programming or teaching. While I could not have told you why, neither of those roles sounded like the right fit for me. Thus, I decided to pursue a master's to qualify for different opportunities.

At the end of that program, I felt educated, but still didn't really understand what one does as a statistician. On the recommendation of my adviser, I entered a PhD program. My adviser thought my academic skills would serve me well in academia, and I realized more education would enable me to be selective when I finally left school to pursue a career.

I worked several hours per week in a statistical consulting lab during that program, helping students, and even did some freelance work helping local businesses. At that point, I started to understand the value I could bring to the business world and the fun of being a statistician.

When I became ready to search for full-time, gainful employment (2004), I had no idea what industry or role I wanted to try first. I must have sent out hundreds of job applications. I used every website I could find and included both the government and public sectors in my search. The only conscious decision I made was to *not* pursue any applications that asked for code samples because, while I had learned a common statistical scripting software for school, I was not comfortable as a coder. I'm not sure that would even be a choice in today's environment, with most entry-level statistician roles requiring SQL, R, python, or experience with some other scripting language.

I accepted a position as a statistician in the consumer insights and product development area at the largest winery in the world after about six months of searching. Oh, my goodness. What a job! I am often asked what a statistician does at a winery. The short answer is they apply everything they have ever learned in school and then learn some more to apply. During the six years I worked at the winery, I primarily supported marketing, production, scientific research, and consumer research, but I think I might have worked with each department in the company at some point.

While working with the marketing department, the statistical applications were straightforward. We designed and analyzed conjoint studies to select new labels or brand names. We also defined control and test markets for point-of-sale marketing and price testing and advertising. We combined our learnings from the control store work with other data purchased from grocery chains and built price elasticity models using ARIMA time series with covariates. None of these methods was explicitly taught in my coursework, but creating a control group follows logically from ANOVA.

My work was predominantly sampling plans for defect testing and the application of statistical product control methodologies in the production environment. I was also involved in some LEAN/Six Sigma projects. I think production work might have been the most 'typical statistician work' I did. In fact, there was so much opportunity for the use of sampling plans and statistical process control methodologies that I developed a training program to educate the process engineers on the methodology so they could run it without my input.

The scientific research was fascinating stuff. The company not only sold wine, but they grew, harvested, and crushed the grapes; made the wine; and aged it. Each step involves nearly infinite variables that can be tuned. Their manipulation affects cost, quality, and the nuanced flavor and aroma of the wine. My role was to work with the scientists to design and analyze their experiments. Some experiments were about multivariate optimization, some about mixtures, and some about simple incomplete blocks. The three experimental design classes I took in graduate school only hinted at the diversity of applications.

The consumer research part of my work at the winery was the original job description; everything else was fun stuff I found to work on. Consumer research was about answering the following questions:

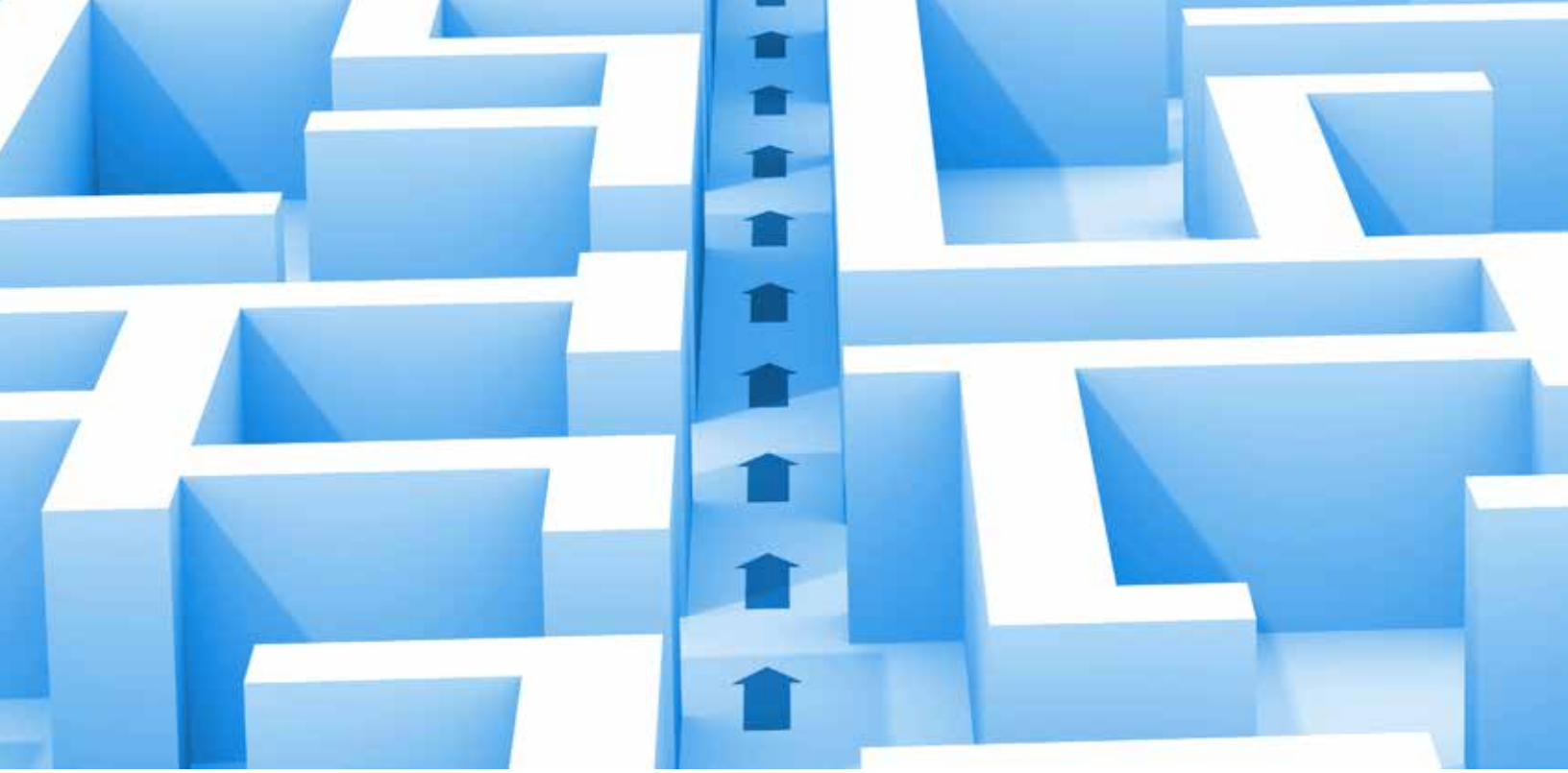
- What kind of wine do people like? *Performing cluster analysis on the wines and people*
- Why do they like it? *Using sensometric and chemometric data with high multicollinearity to model consumer preference measured on a Likert scale*
- Why don't they buy more of it? *Designing and analyzing consumer surveys to discover root cause and potential market impact and recommend product development to meet the consumer need*

After six years at the winery, I chose to explore a different industry. Consumer food and beverage is fascinating, but big data was starting to be a catch phrase so I transitioned to a company focused on internet advertising. The data in that world is definitely big! I joined the team as a statistician, and my role was twofold: create and test targeting algorithms to get the right ad in front of the right people at the right time and analyze recent historical data to identify the root cause of unexpected changes in algorithm performance. In no time, I was leading that team and, shortly thereafter, to managing a larger team. As manager (and eventually director), I was no longer analyzing data, but prioritizing projects and working with business leaders to identify new areas in which the team could add value. The team I managed was made up of statisticians, machine learning scientists, and data engineers. I walked in knowing nothing about computer programming (some C++ in college really doesn't count) and learned enough to help the team succeed.

When I left the winery and went into digital advertising, I started to think about where I wanted to go with my career. I have always said I love the application of statistics because it teaches me about the world. I know so much about how sun exposure in grapes changes the flavor of the wine and what levels of isobutyl methoxypyrazine in combination with low levels of fruity esters will make the wine taste vegetal. I understand the data economy on the internet and what privacy really means. I want to keep learning and be able to apply the numerous methodologies I have learned to more industries. Thus, I decided to try consulting.

I joined a start-up within a giant company in 2014. The company provides products and services to asset managers based on the transfer and processing of data. Given this abundance of data, the company chose to incubate a group that would focus on developing data-driven products and providing data management and data usage consulting to its clients. I came in to build the consulting part of that business. Four years in, I am still building, and the focus has morphed to fit the needs of the market. I spend most of my consulting time helping clients decide where to focus their energy around analytics and advising them about measuring the value of their analytics efforts. In addition to the consulting, because of the exposure I have to the clients, I create solutions for and advise our analytics product development team.

I'm not sure what will come next in my career. Perhaps I will stay on the consulting path. Alternatively, I could go deeper into leading analytics product development. There is even the lure of going back to consumer goods as a leader. I plan to work as long as my brain holds out, and there is no reason I can't explore all these options and probably many more I haven't even dreamed of. ■



A Look Back at My Career as a Biostatistician in Aging and Dementia Research

Chengjie Xiong is professor of biostatistics at Washington University School of Medicine. He earned a bachelor's in mathematics from Xiangtan University, master's in applied mathematics from Peking University, and PhD in statistics from Kansas State University. Read his full bio at *https://bit.ly/2najh3C.*

I started my career as a biostatistician in aging and dementia research by chance. In 2001, after a few years of teaching statistics in a mathematics department, I had to move due to my family's relocation to St. Louis, Missouri.

I joined the research faculty in the division of biostatistics at Washington University School of Medicine in St. Louis to primarily support the research design and statistical analysis of the Knight Alzheimer Disease Research Center (ADRC). I anticipated a challenging transition from teaching to research, given my relatively limited training in biology and medicine, especially in the areas of aging and dementia.

On the other hand, I was attracted to the position on a few fronts.

First, it provided an opportunity for me to work with a huge real-world longitudinal database that could be used to examine the course of Alzheimer's disease (AD) and dementia and aging from a variety of aspects, as well as to be part of a productive and creative research team.

Second, although I had sparse prior experience in biostatistics before taking the position, work in this area aligns with my long-term interest in the applications of statistical theories and models in the biological sciences and appealed to me as an area that could be approached from both methodological and medical/biological perspectives.

Third, many ongoing large longitudinal research projects in AD—both observational and interventional—pose a wealth of challenging and fascinating statistical methodological issues, including the appropriate use of longitudinal linear and nonlinear models, nonlinear mixed model, Markov chain–type of transition models, marginal means and marginal odds ratios models with generalized estimating equations, censored and truncated observations and survival endpoints with competing risks, receiver operating characteristic curves and surfaces, and missing data.

These methodological challenges are, in my opinion, the dream of statisticians who seek to make a difference in real-world applications of statistics.

I was immediately immersed in the extensive interactions between cutting-edge dementia sciences and statistical design and analysis by working on a variety of statistical methodological problems associated with AD studies.

I have sought advice and been fortunate to have had three important mentors: Gerald van Belle from the University of Washington and Philip Miller from Washington University—both experts in statistical applications to aging and dementia research—and John Morris, a leading physician scientist in AD research. I met with these mentors on a regular basis.

As my interest in the statistical methodologies of aging and dementia research grew, I quickly realized I needed to understand and appreciate the clinical and biological aspects of aging and AD. Subsequently, as I gained more exposure to the variety of biomedical ways the natural history and interventions of AD are assessed, I came to appreciate the crucial role biostatistics plays in AD studies and the importance of interaction between AD investigators and biostatisticians. This cultivated my interest in assessing the status of statistical applications and quality of statistical methodologies in AD studies and in identifying possible weaknesses in the use of statistical methods in AD research.

A major driver for me are the many less-than-optimal statistical methods and reports in published AD studies and my desire to help identify the most appropriate analytic methods to deal with statistical challenges arising from AD and aging research. This wish eventually led to a successful application for a Mentored Quantitative Research Career Development Award (K25) from the NIH. The award allowed me protected time to not only receive necessary training and mentoring in biology and aging research, but also to develop ways to improve the quality of statistical applications in AD and aging research. It paved the way for me to grow professionally as a biostatistician by communicating statistical concepts and methods in the languages clinicians and physician scientists speak. Since then, I have gradually assumed more leadership in the biostatistics core facilities that support both the Washington University ADRC and the multi-center international Dominantly Inherited Alzheimer Network (DIAN) and its clinical trials unit.

My continued interactions with a wide range of biomedical investigators in AD and aging have led to several successful applications of independent research grant awards from the NIH to tackle the statistical challenges arising from cutting-edge AD and aging research.

Looking back over the past 17 years as a biostatistician in aging and dementia research, I have fully appreciated that statistics is a vital part of aging and dementia research and that a biostatistician must play a leading role in the design and analysis of these studies to safeguard the integrity of scientific findings. More importantly, to become a leader in a collaborative biomedical research setting, a biostatistician must understand the scientific content and articulate potentially complex statistical designs and methodologies in a relevant scientific context.

Whereas AD has been a high-profile disease for decades in this country and Congress has even passed the National Alzheimer's Project Act to build upon and leverage federal efforts to help change the trajectory of AD and other dementia, I remain surprised about how few statisticians work in aging- and dementia-related statistics. ■

Kreuter Selected as Links Lecturer

The ASA is pleased to announce that Frauke Kreuter—director of the Joint Program in Survey Methodology at the University of Maryland, College Park—has been selected as the inaugural Links Award lecturer for adding important links in the progress of official statistics through leadership in education and training and distinguished contributions to the literature on social and economic measurement.

Barry Johnson, director of the Statistics of Income Division of the IRS, will serve at the "connector" in relating points in Kreuter's lecture to issues facing agencies engaged in the production of official statistics.

The Links Lecture will be held November 5 in Room 100 of the Keck Center at the National Academy of Sciences, 500 5th Street NW in Washington, DC, from 3:00 p.m. to 4:30 p.m. A reception will follow.

The lecture will also be webcast.

For the most current information about the lecture, see *www.amstat.org/ASA/Your-Career/Awards/Links-Lecture-Award.aspx.*

STATS FROM THE ROAD

Rising Star Talks About Statistics Journey, Future of the Profession

Amanda Malloy, ASA Director of Development

I recently had the opportunity to visit with Lucy D'Agostino McGowan, a rising star in the statistics and data science community. Lucy didn't start out on the statistics track, even though it is in her blood—both her father and grandfather are in the statistics field and the ASA community has been part of her life for as long as she can remember. "I grew up with JSM as my family vacation each year," said Lucy.

Lucy completed a dual undergraduate degree in religious and romance language studies with a concentration in Italian. It wasn't until she participated in Boston University's Summer Institute for Research Education in Biostatistics (SIBS) that her interest in the field piqued and she went on to earn a PhD in biostatistics.

Throughout her journey, the ASA community has been an asset for Lucy, inspiring and shaping her as a student and now a postdoctoral fellow at Johns Hopkins. When asked about who her heroes are, Lucy responded:

> Three women who immediately come to mind are Lisa LaVange, Liz Stuart, and Francesca Dominici. They are all incredibly brilliant statisticians and powerful thought leaders in our field. Of course, my heroes also include my father and grandfather, who both inspired my entry into statistics and biostatistics. They are both extremely hard working, excellent communicators and kind and loving people—all qualities I hope to emulate as I pursue this area of work.

Lucy is flanked by her grandfather and father, both of whom inspired her to enter into statistics and biostatistics.

Looking into the future of statistics, Lucy says:

I believe we are moving into an age of information overload, which has three consequences for statisticians. The first consequence is thrilling. When I read about a study now, I want to know exactly where the data came from, how it was collected, and how it was analyzed. Often, if it's something I think is important to me, I want to be able to reproduce the result. Because we can share information so easily now, there is a craving for even more information, even though we're in this space of information overload. This collaboration and eagerness to learn leads to awesome innovations and has been speeding up the rate of advances. For statisticians, this means we are being pushed to innovate and develop best practices for reproducibility. This age of innovation is so exciting and such a great opportunity for us as statisticians.

The second consequence is, with all this information and data, there are constantly new opportunities for developing innovative statistical methods, particularly in the observational research space. Much of the data constantly streaming in are observational by nature and therefore subjected to biases such as selection bias and unmeasured confounding. This is a rich area of statistical methodology, and I believe it is only getting better as we are forced to develop better methods that are scalable to these large data sets.

The third consequence of this information overload is that it can be difficult for the public, or even for data-savvy people, to sift through the signal and noise. I think this means we need to really brush up on our communication skills. We need to be able to adequately distill complex ideas into digestible sound bites. There is a really important balance when doing this between watering down a result and overloading the public with unnecessary information, and as statisticians, this is really the tightrope we are walking now. I think we need to learn the best ways to communicate ideas that are both true and correct, but also condensed and understandable. We have a lot of resources for learning how to utilize a complex statistical method, but we need to be sure to also emphasize the importance of being able to actually explain these methods in a way that people can actually use. ■

Cultivating Future Statisticians on ASA Giving Day

Eager to get involved and become part of the statistics community, Lucy co-founded the Vanderbilt ASA Student Chapter. "Vanderbilt was a relatively new biostatistics program (at that point in its fifth year), so we were looking for ways to get involved in the wider statistics community," said Lucy.

"Fostering relationships is so important to me," she continued. "There is so much we can learn from each other, and this seemed like a great way to dive in and gain access to the brilliant minds involved with the ASA."

Thanks to the Caucus for Women in Statistics Travel Award, Lucy also attended the ASA Women in Statistics and Data Science conference and participated in a panel of R-Ladies who were discussing improving gender diversity in a male-dominated field. "I love this conference so much and was so grateful to be able to attend," said Lucy.

It is so important to cultivate future statisticians like Lucy. You can do that on ASA Giving Day, October 19, when we come together, learn what's being done, and support the future of statistics and data science. Money raised on ASA Giving Day will go to support early-career statisticians, like Lucy, and encourage talented young people to pursue careers in statistics and data science.

Find out more about ASA Giving Day and sign up for a reminder to give at *www.amstat.org/givingday*.

To keep up with Lucy or get in touch, she can be found on Twitter as @LucyStats.

From Spreadsheets to Headlines: A Day in the Life of Data Journalist **Ryan Struyk**

Photo courtesy of CNN

Editor's Note: This article originally appeared on *This*IsStatistics in January. Read more about the great and unique careers of statisticians at *http://thisisstatistics.org*.

What is a data journalist? If you ask Ryan Struyk, data reporter and mobile producer at CNN Politics, it's someone who can turn a data set into breaking news. Data reporters explain important issues, but instead of using human sources to break the news, they work with data sets.

*This*IsStatistics recently had the opportunity to talk with Ryan about being a data journalist and what that means. Ryan explains how data has become a key skill for journalists and offers a glimpse into what it's like to be a data reporter at CNN.

How did you come to pursue a career in statistics?

For as long as I can remember, I've been obsessed with finding little nuggets that can make a big point—especially if they involved numbers. I especially loved doing this with sports growing up, but remember keeping statistics on everything from the books in my bedroom to the free throws in my driveway. I started college with plans to be a high-school math teacher, but opted to take my bachelor's degree in mathematics, with an emphasis in statistics, into the crazy world of journalism.

Tell us about how you use statistics at CNN.

One big way we use statistics is in our public opinion polling. This requires a unique combination of technical and abstract skills. There's a big difference between knowing which questions we should ask and the wording we should use to ask them—but we do both. This means polling, in many ways, is both an art and a science.

Another key component of my job is making statistics accessible, both to an internal and external audience. Let's face it: Most journalists don't pick their career path because they like math. So, the ability to talk about statistics in everyday language—words you might hear in a television script—becomes key to your success.

In the era of big data and widely available statistics, isn't every journalist a "data reporter" in some capacity?

One hundred percent. If you're just reading a typical news story—something we wouldn't necessarily call "data journalism"—odds are it's got a statistic or two in it. Reporters use statistics as proof to back up major points in stories or fact check arguments from politicians all the time.

It's our job to cut through the spin and present the data in as straightforward and accessible a way as possible. That's often easier said than done, but a solid foundation in statistics is usually the first step to getting it right.

Being able to understand and think critically about the way statistics are used can help us navigate the media, optimize the way we do our jobs, and understand the stores we visit and the websites we browse. Especially in today's world of big data, every organization has a huge opportunity to harness the power of statistics.

What do you say to people who say statistics is "boring"?

If you think statistics is boring, you probably don't have a full understanding of what statistics is.

Statistics can save lives, win championships, and solve societal problems. Statistics can win elections, make profits, and cure diseases. For every sweeping goal that aims to change the world, statistics are almost always involved in the efforts to accomplish it. Statisticians have the tools to explain and shape the world on a colossal scale. That doesn't sound too boring to me.

How can statistics help combat the notion of "fake news"?

People with agendas on all sides of all issues can manipulate data to back up just about any point they want to make. But ultimately, the facts matter. In my time as a reporter, I've used statistics to correct the record and fact check everything from the health care debate in the Senate to the Idaho state legislature's efforts to confront a public defender shortage. A robust knowledge of what does (and doesn't) look fishy in the methodology behind real-life statistics is key to defending the facts in this business.

What advice would you give to aspiring data journalists when it comes to working with data?

Keep one foot in both worlds. Understand the nuts and bolts of statistics as well as anyone. Make sure you can hold your own with a big data set and you learn how to make your spreadsheet sing. But then learn what it means to write an engaging television script. Learn the science of snappy headline writing. Know what makes an interview question really sparkle. A full understanding of both the fast-paced, engaging media world and the academic, technical statistics world will serve you well. ■

PASTIMES OF STATISTICIANS

What Does Maura Stokes Like to Do When She Is Not Being a Statistician?

Stokes signs her novel, *Fadeaway*.

Stokes

What do you do in statistics?

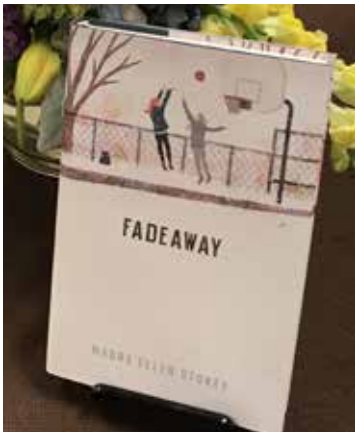
Currently, I am senior director for statistical development at SAS. I also oversee the documentation process for my division and its web presence. I have held various positions at SAS over the years, and prior to that I had a stint at Research Triangle Institute. I have also participated with the ASA organization, serving COPSS [Committee of Presidents of Statistical Societies] and COM [Committee on Meetings] and chairing the ad hoc group charged with beginning the Conference on Statistical Practice.

What are your hobbies?

I ride a bike, play golf, and read.

And?

I write. I consider writing to be my avocation. It's way too much work to consider it a hobby! I also consider it my early morning career, because I do most of my writing from 5 to 7 and on Saturday mornings. Occasionally, I get away for a writing retreat for a few days and just focus on writing.



I just published my first upper-middle-grade novel, *Fadeaway*.

How long has this been going on?

I've been writing since I was a kid. I took a creative writing course in college, and I wrote a couple of columns for the school newspaper. Grad school in biostatistics and getting started in my statistical career limited my time for writing, but eventually I got back to it and earned an MFA in writing at the Vermont College of Fine Arts, with a focus on writing for children and young adults.

So you are really using both sides of your brain?

Yes and no. I'm left-handed, and that means the left- and right-side brain paradigm gets pretty scrambled! Some people ask whether the reasoning and logic of my mathematics background helps with the plotting aspects of writing fiction, and the answer is not at all. Either that or I am so weak at it that the math levels the playing field for me!

So tell us about this book you published in June.

Fadeaway is the story of a 14-year-old girl who loses her best friend right before their freshman year of high school. They are basketball players, and Sam is devastated at the idea of even playing ball again without Reagan. However, Reagan shows back up, at least to Sam, as she navigates the way to life without her best friend. So, it's a story about dealing with grief, but it does have a comic tone to it.

Is the book autobiographical?

Only in that it's set in a small town in New Hampshire, where I grew up, and I did play basketball.

How do you get a fiction book published?

It's not easy! These days, you really need to have an agent, although you can still directly submit to editors—usually those you meet at conferences and workshops and the like. I met an agent through an alumni conference at my school, and she was interested enough in the beginning chapters of this book to keep in touch. Almost a year later, she offered me agency based on the completed manuscript and, after a series of submissions, sold the book to a publisher starting a new middle-grade imprint.

You then go through an editing process, the cover art design, the decision on the title, and—a year or two later—you get a box of your books in the mail! My local bookstore, Flyleaf Books of Chapel Hill, was kind enough to host a book launch for *Fadeaway*. I did a short presentation, read a few sections, and signed books. It was a lot of fun!

Where can you get a copy?

The usual places. You could order it at your book store or online. However, I don't see Amazon offering to bundle it with *Categorical Data Analysis Using SAS*—my other book! ■

STATS4GOOD

Data for Good Goes to College

With the new academic year, students at colleges and universities have a huge number of activities from which to choose, offering an endless supply of experiences and opportunities. There are many ways to connect and make a difference for those seeking to apply their developing analytic skills to projects for the greater good.

While trying to launch a new student group can face challenges—from finding a faculty sponsor to navigating a complex approval process—it isn't necessary to start a whole new student organization from the ground up. Many existing groups engaged in community involvement, social justice, or good causes can benefit from students' work in analytics.

Professionals wanting to get started in Data for Good activities are often advised to look at organizations with which they already have a relationship. The same is true for students. Take a look at the student organization, community service group, fraternity or sorority, sports league, animal shelter, or another place where you are helping now. Sit down with the organization's leaders who already know you and your contributions and have a talk about how you can do even more good with your statistics and data science skills.

Data for Good efforts can be combined with academic work in class projects, term papers, and summer internships and research opportunities. A number of colleges now have entire programs based on using analytics for the greater good. One of the most prominent is the Data Science for Social Good program at The University of Chicago. This renowned program offers summer research fellowships in a variety of analytic skills areas—including machine learning and big data methods—on important Data for Good projects.

Partnerships with community organizations and mentoring relationships develop students' scientific skills and help them become community leaders using statistics and data science to address today's most important issues. Faculty and administrators at other colleges might want to learn more about the Data Science for Social Good program and others to design a program for their own college or university.

One type of student group we sometimes find is a subset of a larger organization, representing the intersection of data analysis and work for the greater good. The Data for Public Good group within the University of Michigan School of Social Work is an outstanding example. Within some 700 graduate students in social work, about 40 are involved in the Data for Good organization, and many more use data analysis in their projects.

I had the opportunity to meet these students and see their great work first-hand when giving the keynote at their annual symposium in April. Instead of the usual conference format with a few full presentations, the Michigan symposium had a large number of teams give a short presentation on their work to the whole group, followed by a poster session. This worked well for a fairly small group deeply involved in statistical projects from within a larger organization in social work. The same could be done for so many areas, such as a public health department hosting a data analytics symposium or a math, statistics, or biostatistics department forming a group to focus on Data for Good projects.

Of course, Data for Good isn't just for students. Faculty, staff members, administrators, student mentors, parents, and even local community organizations all can play a role in supporting Data for Good activities at colleges and universities—recruiting student participants, connecting students with projects and organizations in need of help, supporting and encouraging student work, and creating scholarships and work co-op programs. Local community organizations will want to develop relationships with statistics departments and data science programs, creating opportunities for service and matching students wanting to serve with opportunities in the community where the need is greatest.

With participation in Data for Good, college can be more than making the grade: You can make a difference!

With late summer and fall comes hurricane season. For new Data for Good opportunities this month, consider taking a look at FEMA, especially its Disaster Crowdsourcing Exchange (*https://bit.ly/2n1XXgH*). FEMA recruits volunteers with data skills to help respond to disasters, producing tools such as databases and maps to support relief agencies on the ground. FEMA data volunteers can make a huge difference to people when it is needed most! ■

With a PhD in statistical astrophysics, **David Corliss** works in analytics architecture at Ford Motor Company while continuing astrophysics research on the side. He is the founder of Peace-Work, a volunteer cooperative of statisticians and data scientists providing analytic support for charitable groups and applying statistical methods to issue-driven advocacy in poverty, education, and social justice.

STAT*tr@k*

Making the Meandering Path Go Where You Want It To

Had I ever have thought I would be professor of mental health in a school of public health? Had I ever even heard of that type of job? No! But is it just the right fit for me (despite having no formal training in either mental health or public health and taking a rather circuitous route to get here)? Yes! Hopefully my path will serve as an interesting example of finding your way into the job that's right for you, even if the path isn't clear when you start out.

The graphic illustrates some of the key institutions and decision points in my life, and here is a listing of some of the lessons I learned along the way:

Don't pay too much attention to job labels. I never would have thought to apply for a position as an assistant professor of mental health, but talking to people and learning more about the job made me realize it was a great fit. Likewise, people had told me to never do a joint appointment because of concerns about having two bosses. But my joint appointment has worked out well, in part because the mental health appointment is really "primary," and so promotion decisions, etc., flow through mental health, though I have strong ties (including teaching responsibilities) with biostatistics.

Don't make assumptions about jobs. The jump to Johns Hopkins was scary—it seemed (from the outside) like an intense place, and people had warned me about a "soft money" environment (like Hopkins), where faculty are expected to bring in a substantial portion of their salary in grants. But what those voices often leave out is that this setting leads to a highly entrepreneurial environment, with the ability to spend most of my time on research. And even at a "hard money" institution, I might end up wanting to have substantial grant support to cover summer salary and some teaching buy-out! So, in the end, the "hard money" vs. "soft money" distinction is less strict than it may seem.

As Don Rubin told me, "Don't turn down a job before you know what it is." It takes effort, but it's worth talking to people at each institution to figure out what the place is really like. Get advice from lots of people, since you have to figure out what makes sense for you, not them, and don't be shy about reaching out—in my experience, people are often quite happy to give advice and "pay it back" from their own process and deliberations.

Make the job what you want it to be. During graduate school, I sought opportunities to work on social science and public policy application (even if many of the people around me were doing genetics). At MPR [Mathematica Policy Research], I carved out time for methodological work and teaching on the side. At Hopkins, I have spent my time on the type of applied work I want to do; by establishing myself as the "go to" methods person in the department of mental health, I established my career doing relevant, applied work.

Find good people to work with and network. To me, any work is more fun and rewarding when done with kind, collegial, and smart coworkers. I have been lucky to have that throughout my career and have sought out those environments. I then made sure to network with them and others to help keep up those connections. Personal connections are often underappreciated in workplaces.

Find the right work/life balance for you. I have built my career at the same time as having a family—I have two elementary school–age kids now. Yes, I might work on weekends sometimes (e.g., writing this), but I often work short days during the week or take off early on Fridays to pick the kids up from school. To me, what matters more than the number of hours I work is how many of those hours are spent doing things I don't want to do. I am lucky to have very few of those hours.

Take the leap. I was incredibly nervous to move from my 'safe' job at MPR, where I was happy, but it has been worth the risk to try out academia. While it can be hard to leave an environment with people you like working with, it's worth taking the leap if you're excited about the new possibilities, and our field is small enough that you often find new ways to work with former colleagues. In my case, I still get to consult on MPR projects, which I love.

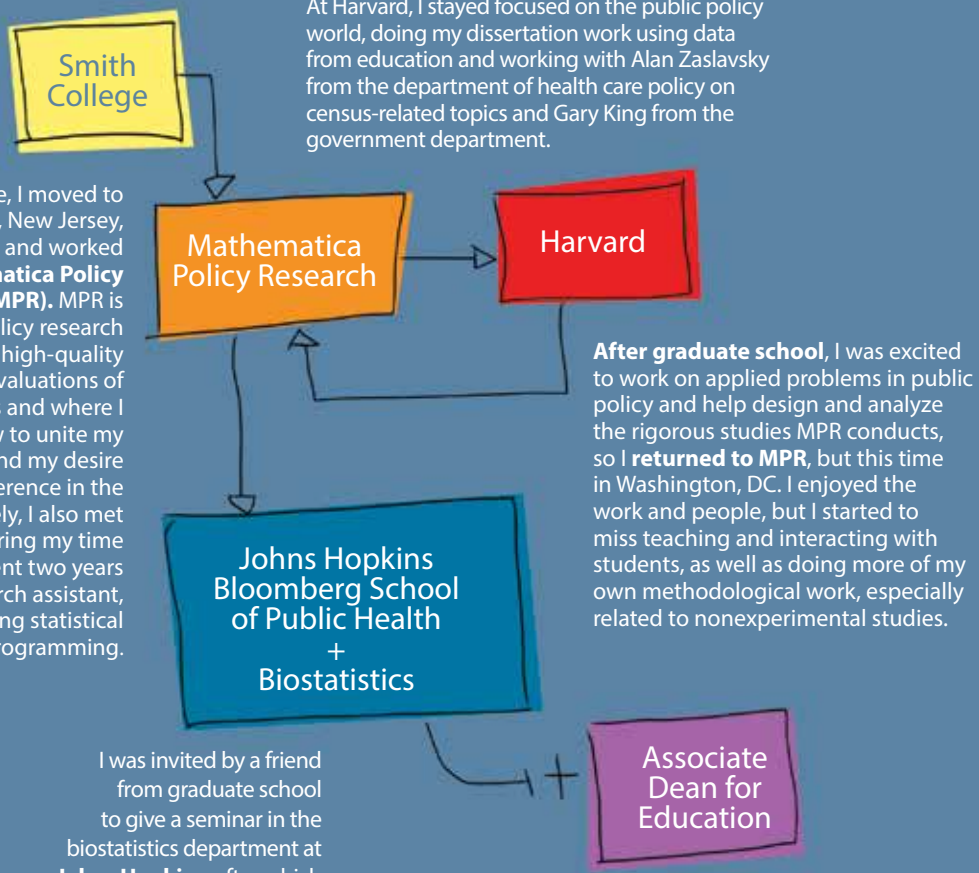
In sum, I have found statistics to be a great field to work in, giving me the ability to work across a range of areas (public health, public policy, education) with great people and interesting problems. I do feel like I'm doing my small part to save the world through math! ■

Elizabeth A. Stuart (*estuart@jhu.edu*) is professor in the departments of mental health, biostatistics, and health policy and management at the Johns Hopkins Bloomberg School of Public Health and associate dean for education. She is a Fellow of the American Statistical Association and has received the mid-career award from the Health Policy Statistics Section, the Gertrude M. Cox Award, and the 2018 Myrto Lefkopoulou Distinguished Lectureship. Stuart has published influential papers about propensity scores and methods to assess the generalizability of randomized trials. For a longer and oral version of this essay, visit *The Effort Report* at *http://effortreport.libsyn.com/47-ignore-the-job-description-a-case-study*.

My undergraduate degree is from **Smith College,** a small liberal arts college. I was a math major and chemistry minor who took a lot of computer science courses. I knew I wanted to apply that mathematical thinking to 'save the world' in some way (lofty thinking … isn't that what undergrads do?), but I had no idea how to do that, never really having heard of statistics or public health at that time.

I knew I wanted to go back to graduate school, so I asked my senior colleagues at MPR what kind of degree would help me get back to doing the type of work being done there. Statistics emerged as the best choice for me. I considered many programs around the country. After many conversations, I ended up at **Harvard,** which gave me my preferred geographic location and the content areas I was hoping to study. In particular, I was excited to work with Don Rubin, who I knew was one of the people who straddled statistics and public policy.

At Harvard, I stayed focused on the public policy world, doing my dissertation work using data from education and working with Alan Zaslavsky from the department of health care policy on census-related topics and Gary King from the government department.



After college, I moved to Princeton, New Jersey, for two years and worked at **Mathematica Policy Research (MPR).** MPR is a public policy research firm that does high-quality rigorous evaluations of social programs and where I discovered how to unite my interest in math and my desire to make a difference in the world. (And, nicely, I also met my husband during my time in Princeton). I spent two years at MPR as a research assistant, mostly doing statistical programming.

After graduate school, I was excited to work on applied problems in public policy and help design and analyze the rigorous studies MPR conducts, so I **returned to MPR**, but this time in Washington, DC. I enjoyed the work and people, but I started to miss teaching and interacting with students, as well as doing more of my own methodological work, especially related to nonexperimental studies.

I was invited by a friend from graduate school to give a seminar in the biostatistics department at **Johns Hopkins**, after which someone suggested I apply for a job in the department of mental health. After a lot of deliberation, I made the leap and left MPR to go to Hopkins in 2006.

I have now been at Hopkins for almost 13 years, and it has been a great experience. I love being able to work on what excites me and basically make the job what I want it to be. I have a strong joint appointment with biostatistics and—given increasing interest in health policy (which brings me back to my original desires)—a joint appointment in health policy and management. About two years ago, I also moved into the dean's office part time as **associate dean for education** and help the school take advantage of data and research to best inform its educational enterprise.

Q&A

BOB KOMARA
Data Scientist at MSX International

What or who inspired you to be a statistician/ data scientist?

I was well on my way to a PhD in physics with undergraduate degrees in math and physics when life got in the way. I realized I did not want to be an academic and many others were leaving academia for high-paying industry jobs. So, after completing a master's and earning the infamous ABD title (all but dissertation), I moved to a Colorado ski town and started a new career managing high-end construction projects. I thought I would be able to bring my scientific background to bear on commercial problems and, to some extent, this was successful. But it wasn't enough to keep me engaged. I kept reading new research and was very interested in the beginnings of the replication crisis and poor scientific reporting/communication being brought to light by different skeptic communities, and this was a distraction.

After a second major career detour that involved working at an aquarium store (I had one of the most beautiful coral reefs in Colorado in my bedroom); being a fly-fishing guide in and around Rocky Mountain Park, Estes Colorado; and launching and managing a state chapter of a national charity aimed at connecting children and their parents back to nature and each other through fishing called Fishing's Future, I met my future wife (an academic psychologist) and we started a budding romance that involved an inordinate amount of heated discussions about data analysis, statistics, linear models, etc. She inspired me to go back to school and get a master's degree in statistics so I could turn my "armchair" data analysis critiques into something more productive and, hopefully, lucrative. It didn't hurt that she was awarded an important postdoc in Switzerland and I needed an official reason for residency to be able to live abroad with her. An incredible bonus was that while most things in Switzerland are expensive, graduate school tuition only cost me about 865 CHF per semester, which was an incredible bargain at about $930.

How did you end up in your current position?

After three beautiful years living in Switzerland, a master's degree, and a marriage, we moved to the UK where I participated in a Knowledge Transfer Partnership (KTP) program for my first job as a data scientist after graduation. A KTP is a joint effort between the UK government, a university, and a commercial business that aims to link academic and business progress to spur innovation. I successfully applied my master's education in spatial statistics/ Gaussian Process regressions and R training with the help of academics at the University of Essex at Objective IT (a database and software development company). Together, we built a successful data science capability in the company by delivering successful products like including classifiers for highly imbalanced data sets and product recommenders for Europe-wide clients. After the successful and award-winning (Best KTP Associate University Essex 2017) two-year project completed, I accepted a new position at MSX International to apply machine learning techniques to complex and interesting problems in the automotive warranty industry.

Name a few specific skills you need to do your job.

- Statistics
- Communication
- SQL
- R
- Vision to see the end goal

Data scientists must provide value. Many people rightfully include it as one of the big data "Vs" now.

What skill would you like to learn to be better at your job?

I am a believer that if you aren't analyzing images, text, or video, then the business value of training interpretable machine learning models with good feature engineering exceeds the benefits of deep learning (for now), but want to take advantage of the power of LSTMs for time series forecasting.

What is the most exciting part of your job?

Getting good test results back on a deployed model and having everyone on the team celebrate the shared success!

What career advice would you give your younger self?

I wish I would have spent more time exploring options and taking risks earlier. I kept pursuing physics because I was good at it, but not necessarily because it was my passion. A better understanding of the exploration/exploitation dilemma would have given me better focus.

Name one or two favorite blogs or books you have read and would recommend to others.

I love reading R-bloggers: *www.r-bloggers.com*. It is an aggregator of several hundred R blogs, and I love digging into tutorials and learning about new R packages featured there.

Two formative books I must mention are *Innumeracy: Mathematical Illiteracy and Its Consequences* by John Allen Paulos and *The Demon-Haunted World: Science as a Candle in the Dark* by Carl Sagan. Both were massively influential in blending my mathematical abilities with critical thinking skills and bringing them to bear on the question of data analysis in particular and science communication in general. ■

What or who inspired you to be a statistician/data scientist?

My grandmother inspired me to become a doctor (MD) when she was diagnosed with breast cancer while I was very young. I promised I would learn how to 'fix' her, but she passed away much too soon for me to fully execute my plan. I had always loved math, so in my mind, when I discovered biostatistics, it seemed the perfect compromise! I could be an 'applied mathematics' doctor instead!

How did you end up in your current position?

After spending 15+ years at the University of Pennsylvania—my first position after graduate school—I was looking to take on new challenges in a new position just after the school of public health at Drexel had gotten a new dean (Ana Diez Roux) and chair of the department of epidemiology and biostatistics (Leslie McClure). Drexel was also planning to launch a new PhD program in biostatistics, which seemed like a great opportunity to pursue, so I did! Both Penn and Drexel have been great fits for me at the respective stages of my career. Wouldn't change much about my career path.

Name a few specific skills you need to do your job.

Besides the obvious (strong statistical knowledge and knowledge of statistical computing software), clear and confident communication skills. And as you take on more administrative responsibilities, leadership skills are also essential. Keeping on top of the literature and [having a] willingness to learn new things is also critical.

What skill would you like to learn to be better at your job?

I would like to learn more about visualization and exploring interesting data science/

SCARLETT L. BELLAMY
Professor and Director of Graduate Studies in Biostatistics, Drexel University Dornsife School of Public Health

statistics intersections. Also, as I am taking on more administrative responsibilities, learning strategies to make me an effective leader is also becoming increasingly important.

What is the most exciting part of your job?

Every day is a new challenge, which is exciting! I love the autonomy and freedom to pick and choose problems I find interesting to work on. I equally enjoy the many ways I get to interact with students at all stages. They have so much energy and lots of great ideas!

What career advice would you give your younger self?

Take risks; confidence work is constant; and there are many ways to be great and yours/mine may or may not look like anyone else's!

Name one or two favorite blogs or books you have read and would recommend to others.

Most of my reading is strictly for fun! I have very strict work/no-work boundaries. I recently finished *We Were Eight Years in Power: An American Tragedy* by Ta-Nehisi Coates and am currently reading *The Awkward Thoughts of W. Kamau Bell* by W. Kamau Bell. I should also plug my chair's blog, StatGirl, which I also follow! ■

Q&A

JULIE NOVAK
Senior Data Scientist, Netflix

What or who inspired you to be a statistician/data scientist?

Funny enough, an *Economist* article. I was an undergraduate student at McGill University, where I was studying theoretical and applied mathematics. I was nearing the end of my bachelor's degree and thinking about next steps. All I knew at that time was I enjoyed mathematics and problem-solving. I have fond memories of brainstorming solutions to challenging problems with friends.

The summer before my senior year, I came across an article from *The Economist*. It predicted that, with technological advances, massive data will be collected across all companies. There will be great demand to analyze and make sense of this data. The opportunities will be endless and exciting. Statistics will lead to unimaginable discoveries and drive decision-making in medicine, marketing, finance, tech, etc. The article also mentioned the work would be highly collaborative, which was another appeal for me. I am very social, so a job where I would have to frequently work with others was ideal. This role would also give me an opportunity to continue learning about new areas throughout my career. I then applied to PhD programs in statistics and went to the University of Pennsylvania to pursue my graduate degree.

How did you end up in your current position?

By chance! Prior to joining Netflix, I was a research staff member at IBM Research and not looking for a new job. This was the first company I worked for after I finished my PhD, and I was fortunate to have a brilliant mentor and manager, Yasuo Amemiya, throughout the two and a half years I worked there.

My current director, Nirmal Govind (who I did not know at the time), sent a note to my personal Gmail, asking whether I would like to have a conversation about data science at Netflix. I was flattered and curious about Netflix, since it is known to have one of the strongest data science departments in the tech industry.

Nirmal told me about the problems his team works on. He also sent me some Netflix Tech Blog posts (which I highly encourage those interested in learning more about Netflix to read). These articles gave me a clear understanding of what the team does. One of their focus areas is improving the quality of the streaming video service to its 100+ million members in more than 190 countries. They do so by building machine learning and statistical models, as well as running A/B tests at massive scale.

As I was going through the interview process, I was consistently impressed with how technically strong and personable my interviewers were, so when I was given the chance to become a senior data scientist on the team, it was an opportunity I could not miss. I currently develop nonparametric statistical methodology to improve the way we analyze A/B test results.

Name a few specific skills you need to do your job.

Besides statistics, the key skills needed for my job are "learnability," coding, and clear communication.

Statistics was the only skill from the list I expected to be important prior to working in industry. Even so, the methodology one learns in school may not be relevant for the work they will do. For example, my PhD thesis was in Bayesian statistics, whereas I focused on forecasting at IBM Research and experimentation at Netflix. This ties in closely with the second skill on the list: learnability. Every time I switched roles or companies, there were required tools I had not used before. For example, when I started as a data scientist at Netflix, I had to upskill on Python, SQL, Tableau, and R Shiny. Being able to learn new concepts and tools on the job is essential.

As for coding, I need to build, test, and debug models efficiently and effectively. Some of the code I develop is then used by other data scientists and stakeholders in the company,

so it cannot break in production. I often work on projects with others, so the code needs to be understandable by others and reproducible.

Finally, the ability to explain your work to other specialists and nonspecialists is crucial. For stakeholders and colleagues to use your methods and trust your results, they need to understand how you got there. By being a clear communicator, you develop stronger relationships and, in turn, your stakeholders come to you with more problems and rely on your data-driven solutions.

What skill would you like to learn to be better at your job?

I cannot think of skills I need to learn from scratch; however, there are certainly ones I want to improve: coding, presenting, and writing. I did not know how fundamental each would be for statisticians. I wish I had been developing them for a longer time! For example, if you build an incredibly powerful model but others cannot understand what you did, it may lose the impact it deserves.

What is the most exciting part of your job?

I would say there are two: impact and stunning colleagues. It's gratifying to work in a highly data-driven company, where decisions are made based on statistical results. If I have a hypothesis, I can run a large-scale A/B test to see whether it improves the streaming service for our customers. If the results show an improvement in streaming quality, the company will move forward with the change. An equally exciting part of my job is working with brilliant people every day.

What career advice would you give your younger self?

I would worry less about getting started in my career, especially while I was a PhD student. There is plenty of time after graduation to develop as a data scientist. School is only the beginning. Over the course of my career, I will have different data science roles and work at different companies. I should be patient with myself and work slowly toward my longer-term goals. I would also spend more time developing the skills I mentioned above (coding, presenting, and writing).

Name one or two favorite blogs or books you have read and would recommend to others.

I recommend the book *Everybody Lies* by Seth Stephens-Davidowitz (an economist, former Google employee, and *New York Times* writer). The author uses internet data and statistical modeling to try to understand how people think. This book is also a great way to demonstrate to your friends and family how data and statistics can help us answer questions about humanity in ways we could never do before.

I also enjoy reading Andrew Gelman's blog, Statistical Modeling, Causal Inference, and Social Science. Gelman is a professor in the department of statistics and department of political science at Columbia University, and I would recommend following his blog. ■

Q&A

LISA LIX
Director of the Data Science Unit in the George and Fay Yee Center for HealthCare Innovation

In honor of this year's JSM being held in Canada, the ASA's Health Policy Statistics Section (HPSS) went to Winnipeg in the great province of Manitoba to interview longtime HPSS member Lisa Lix. Lisa is professor in the department of community health sciences at the University of Manitoba. She is also director of the Data Science Unit in the George and Fay Yee Center for HealthCare Innovation.

The George and Fay Yee Center sounds like an exciting place!

Yes! It is a nifty place, designed to bring transformation to health care in the province. Its eight platforms focus on developing and implementing evidence-based initiatives that improve care and outcomes for all Manitobans. My unit employs about 30 people, including faculty, technical staff, and trainees from biostatistics, bioinformatics, clinical research, and computer science. We develop and analyze data sources around the idea that health care needs to be more patient oriented.

How does your work dovetail with the Canadian health system?

Well, the Canadian health system is not actually universal health care; each province and territory gets an envelope of funding, which is delivered at the provincial and territorial level. What is done is determined by the province. Ultimately, the single-payer health system looks quite different in the different provinces and territories. Unfortunately, this applies to data as well. Hospital records are standardized by the Canadian Institute of Health Information, but physician records and outpatient care databases look different in different provinces. So, there is a strong push for cross-provincial and cross-territorial studies by the Canadian Institutes of Health Research.

It seems like there are interesting funding models based on the partnerships between the provincial governments and health researchers.

This is true. We receive funding from the Canadian Institutes of Health Research's SPOR (Strategy for Patient-Oriented Research). In fact, the Data Science Unit grew out of SPOR funding. SPOR wanted to fund methods hubs in each province, including Manitoba. There is a strong connection between provincial/regional health care decision-makers and research, facilitated by funding models like SPOR.

What's it like living in Winnipeg?

Winnipeg is a fantastic city to live in, very multicultural. The weather may be cold, but the people and culture are not. Twenty percent of the population is indigenous (First Nation and Métis). So, there is much to do in terms of activities that have a multicultural focus. Come visit! ■

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.*

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

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

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

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

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

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

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

California

■ California State University, Fullerton - tenure track faculty position in statistics at the assistant professor level beginning August 2019. Applicants must have a solid theoretical and applied background in statistics. A PhD in statistics or related field is required. The successful candidate will contribute to the statistics community at CSUF through teaching, research, professional activities, and service. Apply at *www.mathjobs.org/jobs/jobs/12065*. Application Deadline: October 8, 2018. EEO Employer.

Iowa

■ Luther College seeks a tenure-track assistant professor of statistics to begin August 2019. The candidate should demonstrate a commitment to teaching excellence at an undergraduate, private, liberal arts college. Duties include teaching six courses per year and developing a research program. Review of applications begins September 21, 2018, and continues until the position is filled. Visit *www.luther.edu* for more details. Apply at *https://luthercollege.hiretouch.com*. EOE.

Massachusetts

■ The Department of Biostatistics and Computational Biology (BCB) at the Dana-Farber Cancer Institute seeks an experienced and highly motivated PhD biostatistician to engage collaboratively with investigators on basic science, animal model, and human research activities in multiple areas of adult oncology and HIV disease. PhD and at least 2 years of collaborative experience are required. Prior experience in oncology and/or HIV a plus. Apply: *www.Click2apply.net/9kbsqtpdkzxkvdm5*. EOE.

Michigan

■ The University of Michigan Department of Biostatistics is seeking applicants for two tenure-track faculty positions for fall 2019. Outstanding senior candidates will also be considered. Candidates must have a strong research background with a doctoral degree in biostatistics, statistics, mathematics, the computational sciences or a related field. For details, visit *https://sph.umich.edu/biostat/faculty-research/job_postings.html*. To apply, visit *https://sph.umich.edu/biostat/faculty-search/index.cfm*. Applications from women and minorities are encouraged. EOE/AA.

■ The Survey Research Center (*www.src.isr.umich.edu*) in the Institute for Social Research at the University of Michigan seeks an established leader in survey research to serve as director of survey research operations. Applicants should submit a cover letter and résumé describing their interest and experience at the UM careers portal: *http://careers.umich.edu/job_detail/159640/director_of_survey_research_operations*. For confidential inquiries, please contact Steven Heeringa, associate director of the Survey Research Center, *sheering@umich.edu*. The University of Michigan is an equal opportunity/affirmative action employer and is responsive to the needs of dual career couples. Women and minority candidates are encouraged to apply.

■ The College of Literature, Science, and the Arts at the University of Michigan is seeking applicants to the Collegiate Fellows program. This

two-year fellowship recruits outstanding scholars in all liberal arts fields, including statistics, who have demonstrated commitments to advancing diversity, equity, and inclusion in higher education. Application deadline: October 1, 2018. Application Information: *http://myumi.ch/JYppY*. The University of Michigan is an equal opportunity/affirmative action employer.

New Mexico

■ Tenure-Track Assistant Professor in Applied Statistics, College of Business, New Mexico State University. A successful candidate will be expected to provide innovative and excellent instruction, advise graduate students, conduct and publish research (collaboratively or individually), and participate in typical university, college, and departmental activities. An earned doctorate in statistics or a closely related field is required. Apply online at *https://jobs.nmsu.edu/postings/30319*. Deadline is October 1,

Faculty Positions Available

STATISTICS
TEXAS A&M UNIVERSITY

The Department of Statistics at Texas A&M University anticipates four tenure-track Assistant Professor positions (outstanding starting Associate Professors could also be considered) to begin September 2019. Completion of all requirements for a PhD/DSc degree in Statistics or a related field prior to beginning employment is required. The department encourages persons from all areas of research to apply, and is particularly interested in expertise in the broad area of data science. In addition to outstanding research, the successful candidate will be expected to teach undergraduate and graduate courses and supervise graduate students. Excellent computing facilities are available and highly competitive startup funding is anticipated.

The Department of Statistics has a tradition of outstanding methodological, theoretic, computational, and interdisciplinary research. Current faculty members actively collaborate with colleagues throughout the whole university. To apply, please visit https://academicjobsonline.org/ajo/jobs/11439. The hiring committee will start to review applications in mid-October. Applications will continue to be accepted until the positions are filled.

The Texas A&M System is an Equal Opportunity/Affirmative Action/Veterans/Disability Employer committed to diversity. Texas A&M University, the College of Science and the Department of Statistics are dedicated to the goal of building an inclusive and culturally diverse faculty and staff who are committed to teaching and working in an environment of academic freedom and equality of opportunity.

HARVARD UNIVERSITY – DEPARTMENT OF STATISTICS
TENURE-TRACK PROFESSOR OF STATISTICS

The Department of Statistics invites applications for a tenure-track professor for the 2019-2020 academic year. We seek strong candidates in any field of statistics and probability as well as in any interdisciplinary areas where innovative and principled use of statistics and/or probability is of vital importance. Candidates should have strong doctoral records and demonstrated teaching and research experience or the promise of achieving such distinction.

The appointment is expected to begin on July 1, 2019. Doctorate or terminal degree in Statistics or a related discipline will be required by the start date. The tenure-track professor will be responsible for teaching at the undergraduate and graduate levels.
Please submit the following materials through the ARIeS portal http://academicpositions.harvard.edu/postings/8342

Applications should include a cover letter, *curriculum vitae*, teaching statement (describing teaching approach and philosophy), research statement, and representative publications, if applicable. Applicants should also submit names of and contact information for 3-5 references. Three letters of recommendation are required, and the application is considered complete only when at least three letters have been received.

Application submission by December 1, 2018 will ensure consideration.

We are an equal opportunity employer, and all qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, disability status, protected veteran status, gender identity, sexual orientation, pregnancy and pregnancy-related conditions or any other characteristic protected by law.

2018. New Mexico State University is an equal opportunity and affirmative action employer committed to assembling a diverse, broadly trained faculty and staff.

Virginia

■ The Eastern Virginia Medical School - Sentara Healthcare Analytics and Delivery Science Institute seeks a full-time faculty position (assistant or associate professor #8501901). Responsibilities include, not limited to, providing HADSI research; teaching in School of Health Professions; generating funding; supervising HADSI faculty and staff on HADSI-related services pertaining to research, biostatistical, and epidemiological support. To apply, please visit *evms.edu/careers.* EOE. ■

NC STATE UNIVERSITY

Department of Statistics | Tenure-Track Faculty Positions

The Department of Statistics at North Carolina State University seeks to hire multiple tenure-track faculty. All ranks will be considered. The target start date is August 2019, although an earlier start date of 1/1/2019 is possible.

Applicants with interests and expertise in theoretical or methodological research in any area of statistics or biostatistics will be considered. Candidates with interests in modern methods of data analysis are especially encouraged to apply. The ability and desire to supervise graduate student research and to pursue excellence in teaching are essential.

To apply, please visit *https://jobs.ncsu.edu/postings/102311*.

The Department provides a dynamic environment for teaching, research and collaborations across disciplines. Inclusiveness and diversity are academic imperatives and are university goals. We are particularly interested in candidates who have experience working with students from diverse backgrounds and a demonstrated commitment to improving access to higher education for students from underrepresented groups. The Department's location in the Research Triangle provides rich opportunities for interactions with industry; other universities, including Duke University and the University of North Carolina at Chapel Hill; and government agencies. Faculty enjoy collaborations with medical researchers at Duke, environmental scientists at the EPA research facility, pharmaceutical researchers at Glaxo-SmithKline, and software developers at SAS Institute, among many others. The Department is also a founding cooperator of the NSF-funded Statistical and Applied Mathematical Sciences Institute (SAMSI), located nearby in Research Triangle Park.

All applicants must have a Ph.D. in Statistics or Biostatistics by the time of employment. Review of applications will begin soon, and continue until the positions are filled. Questions about the search may be directed to the Search Committee Chair (*stat_search@stat.ncsu.edu*).

Please upload the names and contact information for your three letters of reference at *http://www.stat.ncsu.edu/references* so we can track the letters received from your references. All reference letters will be made available for review by the search committee.

NC State University is an equal opportunity and affirmative action employer. Women and members of other underrepresented groups are encouraged to apply. In addition, NC State University welcomes all persons without regard to sexual orientation or genetic information.

UCI Department of Statistics

Tenure-Track/Tenured Faculty Positions

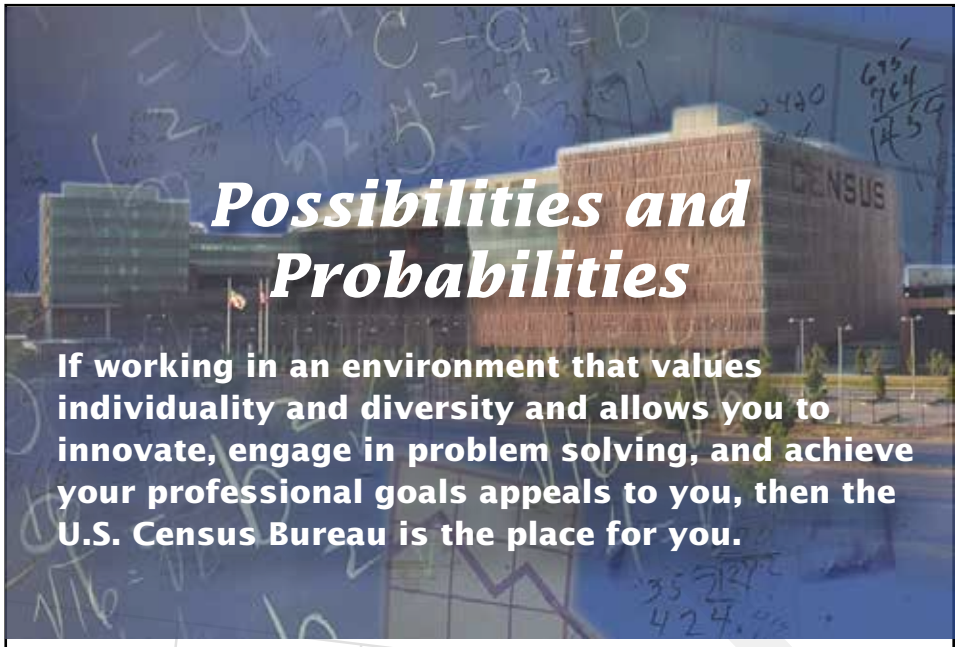
The Department of Statistics (*www.stat.uci.edu*) in the Donald Bren School of Information and Computer Sciences (ICS) at the University of California, Irvine (UCI) invites applications for multiple tenure-track or tenured faculty positions at the assistant or open-rank levels beginning July 1, 2019. Two positions are at the assistant professor level and two positions are open rank.

The Department of Statistics has a strong interdisciplinary flavor, focused on developing state-of-the-art methods for solving topical data-driven problems in science and engineering, and advancing the statistical theory that underlies those methods. We are searching for faculty with strong research potential, a commitment to excellence in teaching, and enthusiasm for helping our collegial department continue to grow. Applicants should hold a Ph.D. degree (or expected by fall 2019) in statistics, biostatistics or a related field. Candidates with research interests in all areas of statistics will be considered.

The University of California, Irvine is part of the premier public university system in the world. UCI is a member of the Association of American Universities (AAU), is ranked as a top ten public university by U.S. News and World Report, and was identified by the New York Times as No. 1 among U.S. universities that do the most for low-income students. UCI is located in Orange County, 4 miles from the Pacific Ocean, 45 miles south of Los Angeles, and 80 miles north of San Diego. Irvine is one of the safest communities in the U.S. and offers a very pleasant year-round climate, numerous recreational and cultural opportunities, and one of the highest-ranked public school systems in the nation. All positions would be eligible to participate in UCI's faculty housing program.

Completed applications containing a cover letter, curriculum vita, graduate transcripts (for assistant professor candidates), statements on diversity, teaching, and research, three letters of recommendation, and sample research publications should be uploaded electronically. Please refer to the following web site for instructions *https://recruit.ap.uci.edu/JPF04670*. Applications received by **November 1, 2018** will receive the fullest consideration.

The University of California, Irvine is an Equal Opportunity/Affirmative Action Employer advancing inclusive excellence. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, sexual orientation, gender identity, national origin, disability, age, protected veteran status, or other protected categories covered by the UC nondiscrimination policy. A recipient of an NSF ADVANCE award for gender equity, UCI is responsive to the needs of dual career couples, supports work-life balance through an array of family-friendly policies, and is dedicated to broadening participation in higher education.



Your Work as a Mathematical Statistician at the Census Bureau

- Design sample surveys and analyze the data collected.
- Design and analyze experiments to improve survey questionnaires and interview procedures.
- Improve statistical methods for modeling and adjustment of seasonal time series.
- Perform research on statistical methodology that will improve the quality and value of the data collected.
- Publish research papers and technical documentation of your work.

Requirements

- U.S. citizenship
- Bachelor's, Master's, or Ph.D with at least 24 semester hours in math and statistics (see Web site for more specifics on required coursework)

Apply at www.census.gov, click on Census Careers, Type of Position, Professional/Scientific/Technical, Math Statistician

The U.S. Census Bureau is an Equal Opportunity Employer.

United States Census Bureau | U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

AMSTATNEWS

ADVERTISING DIRECTORY

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

misc. products and services

professional opportunities

software

SOCIAL CHATTER

What our followers said online during JSM:

FOLLOW US

community.amstat.org

www.facebook.com/AmstatNews

@AmstatNews

www.instagram.com/AmstatNews

This month, we ask: Heard any good math/statistics jokes? Share them with us! Tag @AmstatNews with your answer.

Sherri Rose • @sherrirose
Thrilled to lead @biostatistics as the new Co-Editor starting in 2019!

This will mark a major milestone: I will be the first woman to hold the position of Editor in the history of the journal. #JSM2018

Sean Kross • @seankross
Paraphrasing @mgymrek:

Your paper is cited outside of your field in one semi-related paragraph: *counts towards your career progression*

Your academic software package has 100 stars on GitHub: *counts for nothing*

Something is wrong here #jsm2018

Brian Zaharatos • @bzaharatos
DO NOT DRINK THE TOILET WATER. Whew. Good thing I saw this. #JSM2018

S. A. Santorico, PhD • @sasantorico
Learned about so many opportunities for #DataScience for #socialGood and #dataForGood at #jsm2018 Looking forward to sharing info with students

ASA Science Policy • @ASA_SciPol
The ASA president, prez-elect, & past pres will all be women in 2019, for the first time in history, as @LavangeLisa points out at #JSM2018

Michael Hoffman • @michaelhoffman
It was a real pleasure attending my first JSM. Thanks to the organizers for your work. Enjoyed meeting new colleagues, seeing old friends, and learning the latest in stats research. #JSM2018

Tiffany Timbers • @TiffanyTimbers
LOVE the term coined by Rebecca Nugent: "Behavioural Data Science", and the amazing effort she and others at @CMU_Stats are going to to measure how people do and learn #datascience #jsm2018

Amelia McNamara • @AmeliaMN
Live coding is like a reality TV show– the students are watching and waiting for you to do something embarrassing. #JSM2018

Raja Doake • @doorisajar
Now that I'm sitting in the last room I'll visit during #JSM2018, I decided to add up the walking I've done since I arrived in #Vancouver by looking at my Google Maps timeline. NINETEEN MILES.

Women in Statistics and Data Science

is ...

Welcoming

Inspiring

Empowering

Motivating

Eye Opening

Awesome

... you.

Hotel Reservations Close
September 25

Students and mid-career and experienced professionals all will be represented!

conference
Cincinnati, Ohio
October 18-20, 2018

Learn more and become involved at ***ww2.amstat.org/wsds***.

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.