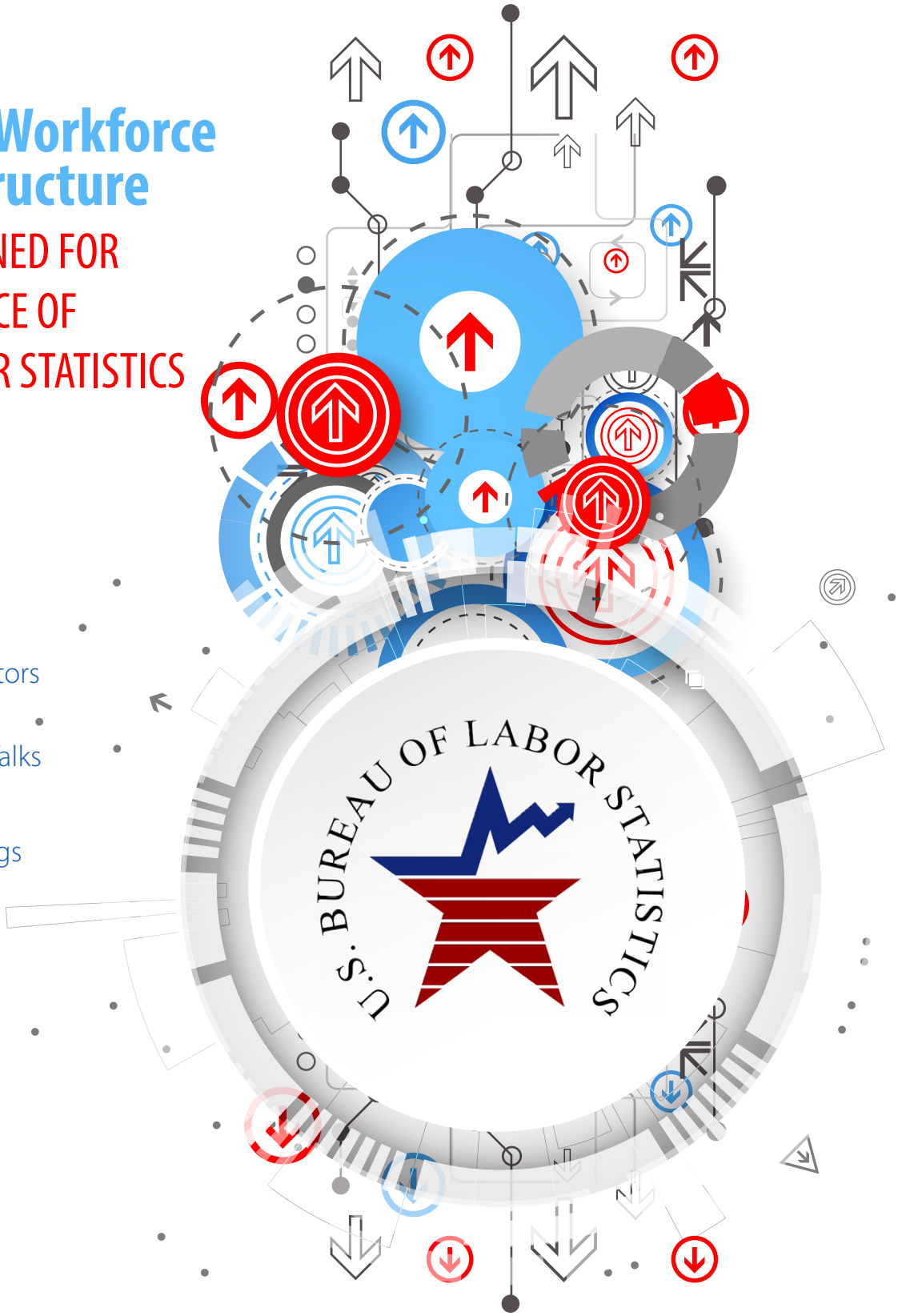
December 2021 • Issue #534

AMSTATNEWS

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

State of the Workforce Data Infrastructure

EXPERTS CONCERNED FOR FUTURE RELEVANCE OF BUREAU OF LABOR STATISTICS



ALSO:

Census Quality Indicators Task Force Wraps Up

Claire McKay Bowen Talks Data Privacy

2022 Internship Listings

CSP

New Orleans, Louisiana
February 1–3 2022

THANK YOU TO OUR SPONSORS

Principal

STATA®

INTERESTED IN SPONSORING CSP 2022?

The American Statistical Association invites sponsors, exhibitors, and advertisers to join us during this gathering of hundreds of statistical practitioners and data scientists. Your commitments accumulate toward your sponsorship status.

CONTACT

AMY FARRIS
Director, Marketing and Membership Development,
American Statistical Association
amy@amstat.org

View the online Opportunities Guide at
ww2.amstat.org/meetings/csp/2022/sponsor.cfm

VISIT *WW2.AMSTAT.ORG/CSP* FOR MORE INFORMATION.

AMSTATNEWS

DECEMBER 2021 • ISSUE #534

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

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

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

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

features

See My ASA Story with Jeri Mulrow

Page 20

Hongyuan Cao, Jeri, and Renee Moore during the 2010 People to People trip to China

columns

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

COVID-19

COVID-19 PANDEMIC RESOURCES

The board and staff of the American Statistical Association have compiled on-demand sources that explain concepts critical to understanding and evaluating the scientific information available to the public. This comprehensive list is online at *www.amstat.org/ASA/Education/COVID-19-Pandemic-Resources.aspx*.

GEORGE MASON WELCOMES EIGHT NEW FACULTY

Mason Statistics (*stat.gmu.edu*) **welcomes eight new faculty: Ben Seiyon Lee**, assistant professor; **David Kepplinger**, assistant professor; **Jonathan Auerbach**, assistant professor; **Lily Wang**, professor; **Mary Meyer**, visiting professor; **Isuru Dassanayake**, assistant professor; **Kenneth Pasiah**, assistant professor; and **Inchi Hu**, professor. GMU is Virginia's largest public research university. The department faculty expertise can be found at *https://statinova.gmu.edu/expertise*. Read more about these new professors at *https://magazine.amstat.org*.

STATS4GOOD

DATA FOR GOOD IN A POST-PANDEMIC WORLD

This month in Stats4Good, we will take a look at the post-pandemic "new normal," where the experiences and technology of working remotely are changing Data for Good—for good! Read STATS4GOOD at *http://magazine.amstat.org*.

IN MEMORIAM Sadly, **Edmund Gehan**, **Bob Hayden**, and **Elliot A. Tanis** passed away recently. To read these members' obituaries, visit *http://magazine.amstat.org*.

departments

INSIDE THE PRINT ISSUE

Look for your 2022 calendar and poster in the center of this issue

member news

Follow us on Twitter
www.twitter.com/AmstatNews

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

Like us on Facebook
www.facebook.com/AmstatNews

Follow us on Instagram
www.instagram.com/AmstatNews

Subscribe to our YouTube channel
www.youtube.com/user/AmstatVideos

The Exquisite Art of Balancing Rigor and Purpose and a Fond Farewell

Photo/Steve Everett

Rob Santos

I once said to a journalist that if I was on a beach and suddenly observed an approaching tsunami, I wouldn't need to know the wave height was 23.9 feet, approaching at 30.7 miles per hour—plus or minus 0.5—at a 95 percent level of confidence. All I'd need to know was a huge wave was approaching fast, so I better get the heck out of there with great haste (paraphrasing and embellishing). In fact, anyone on the beach who took the time to collect precise measurements would likely achieve results too late or not at all for them to make an important decision. Fate would intercede.

Similarly, if I wanted to try a different restaurant in Austin, I would likely consult one of the many online crowd-sourced eatery ratings, knowing full well such ratings are nowhere near a true representation of patron satisfaction and some businesses submit their own ratings to bolster their business or disparage the competition. You just keep that in mind when you see the ratings and take your chances.

At the other extreme, a large-scale clinical trial of a COVID vaccine featuring tens of thousands of participants is needed to demonstrate vaccine efficacy and gauge negative side effects. Here, high levels of scientific rigor and statistical precision are necessary to ensure and advance public health.

Turning to the climate, I rely on well-developed —albeit far-from-certain—hurricane projection models to decide how to prepare for an approaching super-storm. I often marvel at the National Hurricane Center's assorted model projections plotted onto a map to see how congruous or convergent (or squiggly) their trajectories appear. Even when model trajectories vary considerably, sufficient information can be conveyed to allow an effective safety plan to be prepared.

Both sets of examples illustrate uncertainty of various types: systematic error from biases and perturbations from random error. Both also illustrate how we and greater society live with varying levels of uncertainty and quality of information to make decisions affecting our lives. And both exemplify the importance of aligning uncertainty and data quality with our ultimate objectives.

The quality of the information we need is directly tied to the specific task at hand. Experiencing a lousy dinner at a restaurant you chose via potentially unreliable online ratings is admittedly annoying, but it's not the end of the world. Fleeing a beach because you see a large wave approaching can save your life, even though it is based purely on a single, inexpensive, and potentially imprecise personal observation (i.e., n=1). And the result of a large, high-cost, time-consuming, rigorous clinical trial is well worth the investment of resources and time when it comes to advancing public health. This brings us to one of my favorite topics: the concept of "fitness for use."

"Fit for use," aka "fit for purpose," is critical to the statistical and broader research communities. You don't need a multi-million-dollar rigorous statistical survey to understand what population segments prefer tea over coffee and why. Well-designed focus groups may suffice. But rigorous, high-cost statistical surveys are needed to estimate the US unemployment rate or the Consumer Price Index. These statistics can literally affect the global economy and thus need to be of the highest quality.

That is why researchers—including statisticians—must align project timelines, available resources, and methodologies to achieve explicit research objectives. Doing so makes a study "fit" for its intended "use" (i.e., that of achieving the research objective). It takes creativity and technical skill to achieve fitness for use; it is an exquisite art!

As an aside, this reminds me of the age-old iron triangle: "Fast; cheap; good: Pick two." Presumably, it's impossible to achieve all three, although I

We are all part of an amazing statistical community and deserve to help and support each other.

suspect there are some rare examples. Note the COVID vaccines were fast and good, albeit very expensive. But getting back to the concept of fit for use, there are inevitably occasions when scientific rigor, time, and resources/methods simply cannot be aligned to achieve the research objective. In such circumstances, it is always good practice to revisit each element to see if an acceptable revision can be made to accommodate the research objective. But when that is not possible, the research objective will need to be modified. Or the study may need to be terminated at the planning stage.

I long ago stopped counting the instances in which potential clients contacted me to conduct a high-quality survey of an important topic, only to be gobsmacked by the requisite cost and timeline. Fortunately, an in-depth conversation often revealed they did not need a high-quality survey to address their research question. The insights they sought could be gained in a timely fashion with a different, inexpensive research approach.

A great example of this involved the collection of member feedback on personal experiences of sexual assault and harassment at ASA events by the Task Force on Sexual Harassment and Assault. I served as a member of this group. I and others quickly advised against conducting a scientific sample survey because of the almost certain low participation rate and high cost associated with a sensitive-topic study. And, in fact, this was not necessary. It was only necessary to document that harassment/assault experiences befell some of our members, as well as the range of scenarios and manifestations in which these appalling interactions occurred. Those could be collected by simply providing all association members an opportunity for confidential feedback. Many members chose to participate using a feedback form developed by the task force members. Thus, the task force gathered the needed data to garner findings, reach conclusions, and make recommendations.

Being outside of academia for decades, I am unaware of the extent to which the concept of fit for use has found a place in the statistical curriculum. I hope it has, because it is foundational to the design of effective research and knowledge gain. In fact, I'd say invoking the concept of fit for use is one of my secrets to success. Try it out and see for yourself!

In closing, I will note with sadness that this is my final President's Corner column. I hope you enjoyed reading these as much as I enjoyed writing them. I worried I would need to stop earlier in the year due to my nomination for US Census Bureau Director. But I will begin this new chapter in January 2022, at which time I will regrettably resign from the ASA Board.

The opportunity afforded by these columns allowed me to promote what I thought was most important this year—sharing ourselves to build community. We are all part of an amazing statistical community and deserve to help and support each other. Please remember that.

It's been a privilege and honor to serve you as president. I leave you in the superb hands of the board, each member of whom is an exceptional leader. I especially thank Executive Director Ron Wasserstein, Donna LaLonde, and the entire ASA staff for their untiring support. And special thanks to incoming president, Kathy Ensor, who took on additional board duties once my nomination became public. Finally, I thank you, our beloved ASA membership, for your volunteerism (i.e., sections, chapters, outreach groups, caucuses, etc.) and your support of each other during these tumultuous times of COVID. I will always hold fond memories of my time serving you.

ASA Giving Day 2021 Raises Nearly $52,000

Amanda Malloy, ASA Director of Development

For more than 180 years, the ASA community has stood together to embrace opportunities and respond to crises facing the statistical community and society at large. We have learned from and are inspired by those who have come before us—continuing to grow and empower others through knowledge and data.

ASA's fourth annual Giving Day, which took place October 8, celebrated this rich history and a dedication to making a better tomorrow. The ASA community demonstrated this by donating close to $52,000. Thank you to everyone who donated! These donations will help enhance statistics education at all levels, improve the public's understanding of how to make better decisions based on reliable data, advocate for sound policy and governance, and cultivate the next generation of leaders in statistics and data science.

A special video series was launched on ASA Giving Day that featured members discussing issues important to them. If you haven't already, you can still watch the videos at *www.amstat.org/givingday.* Audio files are also available to download and listen to as you would a podcast. Videos in the series include the following:

Building Trust in Federal Statistics

Katherine Wallman – Former Chief Statistician of the United States

Joe Hotz – Arts and Sciences Distinguished Professor of Economics, Duke University

Denice Ross – Senior Fellow, National Conferences on Citizenship

Bringing Statistics to a Wider Audience

Regina Nuzzo – ASA Senior Advisor for Statistics Communication and Media Innovation

Aparna Nathan – 2021 AAAS Mass Media Fellow at *The Philadelphia Inquirer*

Nick Thieme – 2017 AAAS Mass Media Fellow at *Slate*

Innovation in K–12 Statistics Education

Donna LaLonde – ASA Director of Strategic Initiatives and Outreach

Christine Franklin – ASA K–12 Statistical Ambassador

Leticia Perez – GAISE II author and K–12 educator

Fostering Justice, Equity, Diversity, and Inclusion in the Statistical Community

Ron Wasserstein – ASA Executive Director

Adrian Coles – Chair, Committee on Minorities in Statistics, and Co-Chair, Anti Racism Task Force

Kimberly Sellers – Chair, JEDI (Justice, Equity, Diversity, and Inclusion) Outreach Group

Giving Day Chapter and University Challenge Winners

Chapter Challenge Winners
Region 1 – Washington Statistical Society
Region 2 – Ann Arbor Chapter
Region 3 – San Antonio Chapter

University Challenge Winners (tied)
Kansas State University
Northwestern

Malloy

MORE ONLINE
To learn more about the Giving Day challenges and what donations do, visit *www.amstat.org/giving.*

ASA GivesBack Hallow-Meme Contest Winners

The ASA GivesBack leadership team held a Hallow-meme contest in October. Participants were asked to create a statistics meme using statistical concepts and images. Two winners were selected for this inaugural contest: **Larry Lesser** from the open category and **Vasu Prasad** from the K–12 category.

Larry Lesser—Open category

Vasu Prasad—K–12 category

Census Quality Indicators Task Force Wraps Up

Ron Wasserstein, ASA Executive Director

Task Force Members

Nancy A. Potok (co-chair), Chief Statistician of the United States (2017–2020); Former Deputy Director, Census Bureau

Robert Santos (co-chair, resigned in April after being nominated as the next census director), Vice President and Chief Methodologist, Urban Institute; 2021 President, American Statistical Association

Constance F. Citro, Senior Scholar, National Research Council, Committee on National Statistics

Robert E. Fay, Senior Scientist, Westat; Former Senior Mathematical Statistician, Census Bureau

Robert Groves, Provost, Georgetown University; Former Director, Census Bureau

Howard Hogan, Adjunct Statistics Professor, The George Washington University; Former Chief Demographer and Chief of the Decennial Statistical Studies Division, Census Bureau

D. Sunshine Hillygus, Political Science Professor, Duke University; Former member, Census Scientific Advisory Committee

Thomas A. Louis, Biostatistics Professor Emeritus, Johns Hopkins Bloomberg School of Public Health; Former Chief Scientist, Census Bureau

Kenneth Prewitt (resigned in April upon becoming senior adviser in the office of the director of the US Census Bureau), Carnegie Professor of Public Affairs, Columbia University; Former Director, Census Bureau

Denice Ross, Fellow, Beeck Center for Social Impact + Innovation, Georgetown University

Matthew Snipp, Burnett C. and Mildred Finley Wohlford Professor of Humanities and Sciences, Department of Sociology, Stanford University; Director, Institute for Research in the Social Sciences

John H. Thompson, Distinguished Institute Fellow, Biocomplexity Institute and Initiative, University of Virginia; Former Director, Census Bureau

Katherine K. Wallman, Chief Statistician of the United States (1992–2017); 1992 President, American Statistical Association

Steve Pierson, ASA Director of Science Policy (ASA liaison)

Researchers Selected by the Task Force to Analyze Internal Census Data

Jonathan Auerbach, Assistant Professor of Statistics, George Mason University; former ASA Science Policy Fellow

Paul Biemer, Distinguished Fellow, Statistics, RTI International

Joseph Salvo, Senior Adviser, National Conference on Citizenship

After an intense year of biweekly meetings, the 2020 Census Quality Indicators Task Force wrapped up its work in mid-September. Its products were two reports: *2020 Census Quality Indicators: A Report from the American Statistical Association,* released in October 2020, and *Final Report: 2020 Census State Population Totals: A Report from the American Statistical Association Task Force on 2020 Census Quality Indicators,* released in September 2021.

The 13-member task force of census experts was formed last fall by ASA member **Nancy Potok,** former Census Bureau deputy director and former US chief statistician, in response to concerns about the quality of the 2020 Census data, including published opinions by notable ASA members.

For example, four former Census Bureau directors—**Vincent Barraba, Robert Groves, Kenneth Prewitt,** and **John Thompson**—made a public statement last August in which they urged Congress to task an apolitical, trusted institution to "produce predetermined quality metrics that can assess if the final 2020 numbers match other historical and reasonable estimates of the population the Census Bureau produces."

The following month, Potok joined with former US chief statistician and former ASA president **Katherine Wallman** to write an op-ed for *The Washington Post* in which they pointed out that outside expertise had been enlisted in assessing census data in the past and that the National Academy

Further Reading

2020 Census Quality Indicators. A Report from the American Statistical Association: *https://bit.ly/3bT9rgb*

Final Report: 2020 Census State Population Totals: A Report from the American Statistical Association Task Force on 2020 Census Quality Indicators. *https://bit.ly/3o5YaP2*

ASA Board Releases 2020 Census Quality Indicators: *https://bit.ly/2YuXJoV*

On the Importance of Extending the 2020 Census Statutory Deadlines to Achieve a Fair and Accurate Enumeration of the United States *https://bit.ly/3wrvXWU*

Political Interference Threatens the Census. Here's How to Protect It. *https://wapo.st/31KDapM*

Updates on the 2020 Census Quality Indicators: *https://bit.ly/3wl7oVX*

of Sciences and ASA both "have deep experience studying the census that would suit them to this task."

The task force began meeting in September 2020, after the invited members agreed to serve. It was initially co-chaired by Potok and ASA President-elect **Robert Santos,** working under the aegis of the ASA, with biweekly updates on its work provided to the ASA community and broader public. (In April, after President Joe Biden announced he planned to nominate Santos as Census Bureau director, Santos resigned and the task force continued under Potok's leadership.).

The recommendations of the task force's first report, *2020 Census Quality Indicators,* were endorsed by the ASA Board of Directors in October; the task force recommended a series of quality indicators be used by both Census Bureau staff and qualified external researchers to assess the quality of the 2020 Census. In response, the Census Bureau agreed to allow a group of ASA researchers with "special sworn status" to access and apply the quality indicators to confidential, internal 2020 Census data.

The group's final report, *2020 Census State Population Totals,* incorporated information from several input reports: an analysis of the quality indicators using internal census data from **Paul Biemer, Joseph Salvo,** and **Jonathan Auerbach;** a review of research supporting the use of administrative record enumeration in the 2020 Census by **Robert Fay;** and a discussion of demographic benchmarks by **Howard Hogan.**

The task force concluded in this report that, despite concerns about political interference in the 2020 Census, it found "no evidence of anything other than an independent and professional enumeration process" and "no major anomalies that would indicate census numbers are not fit for use for purposes of apportionment." It also recommended the Census Bureau and National Academy of Sciences Committee on National Statistics examine census data at more detailed levels of geography and planning for the 2030 Census should include explicit attention to evaluating and reporting on data quality in a timely manner.

As the task force wraps up its work, the National Academies Committee on National Statistics' Panel to Evaluate the Quality of the 2020 Census is already ramping up. Two task force members, **Thomas A. Louis** and **Matthew Snipp,** are serving on this panel, which will review and evaluate the quality of the data collected in the 2020 Census. Their first report is expected to be released in March 2022 and their final report in May 2023. ■

NOMINATE Colleague for ASA LEADERSHIP POSITION

Nominations are being sought for ASA president-elect and vice president candidates for the 2023 election. While the 2022 elections have yet to be held, the Committee on Nominations needs time to evaluate recommendations to propose the best possible slate of candidates for these critical positions.

As a member of the ASA, you recognize the importance of leadership in our diverse, complex, and multidisciplinary field. You and all fellow ASA members deserve visionary leaders who can ensure our discipline has a voice at the table when appropriate, whether it be in academe; research firms; federal, state, or local government; or nonprofit organizations. This is why we need your input.

For this election cycle, the president-elect will be selected from academe and the vice president will be selected from industry. Think about your colleagues and associates who are members of the ASA and would make good candidates for these positions. Think about members who have helped run a conference or are active in your section or chapter. Then, nominate your choices for the 2024 president-elect and vice president by email to *elections@amstat.org.*

In addition, for the first time, the position of ASA treasurer will be an elected position. Consider nominating individuals for this key board leadership role, as well.

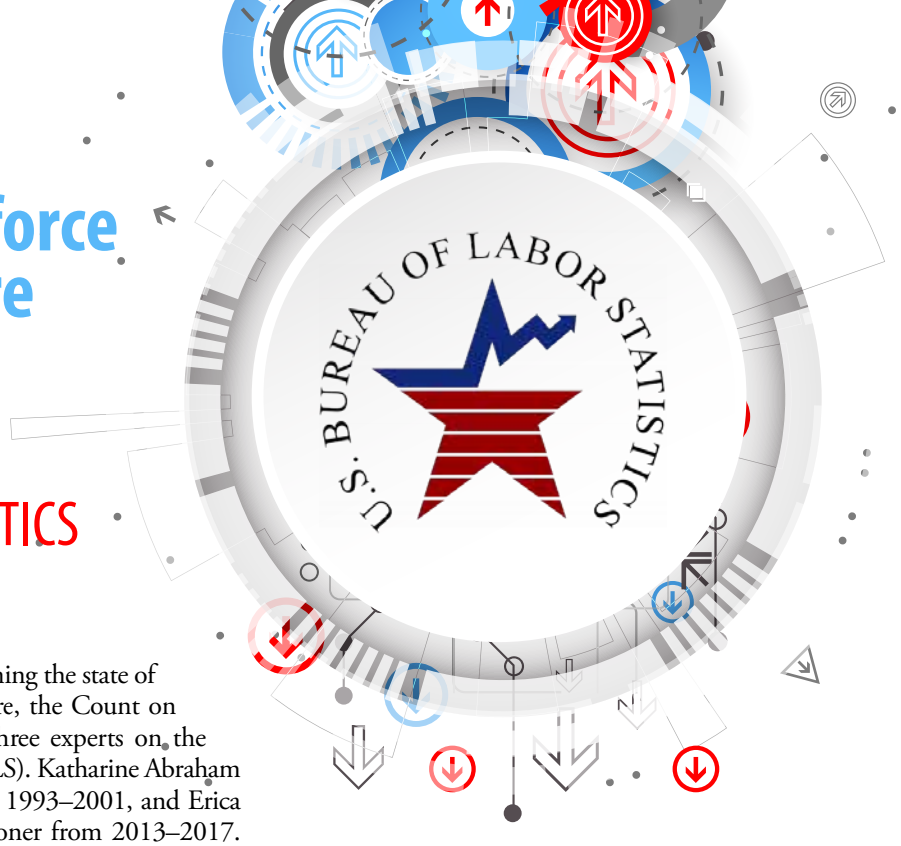
Supply as much information about your nominee as possible to assist the committee in researching them thoroughly and discretely.

The deadline for nominations is February 1, 2022. ■

State of the Workforce Data Infrastructure

EXPERTS CONCERNED FOR FUTURE RELEVANCE OF BUREAU OF LABOR STATISTICS

Steve Pierson, ASA Director of Science Policy



In the latest feature examining the state of the US data infrastructure, the Count on Stats team spoke with three experts on the Bureau of Labor Statistics (BLS). Katharine Abraham was BLS commissioner from 1993–2001, and Erica Groshen was BLS commissioner from 2013–2017. Ken Poole is the chief executive officer of the Center for Regional Economic Competitiveness and long-time and widely recognized leader in the BLS data user community.

BLS is the principal federal statistical agency measuring labor market activity, working conditions, price changes, and productivity in our nation's economy. Established in 1884 as the Bureau of Labor in the Department of Interior, and after 15 years as an independent department starting in 1888, it was incorporated into the Department of Commerce and Labor in 1903. Ten years later, it was transferred into its current host agency, the Department of Labor, upon the department's creation.

The second-largest federal statistical agency by budget, the BLS is also among the few such agencies with the highest stature and strongest autonomy. It is one of only two of the 13 principal federal statistical agencies reporting to the office of the host agency head and one of only three whose leader is presidentially appointed and senate confirmed—for a fixed term, rather than serving at the pleasure of the president. BLS also has strong control over its budget—once determined by the administration and congress—and publications. For information technology (IT), there is no interference with such operations to collect, process, and disseminate statistical information, though there has been repeated interest in incorporating BLS IT into the Department of Labor.

Despite such strengths, the experts interviewed here express deep concern about the challenges the agency faces to meet data user demands for real-time, granular data amidst growing competition from private sector data providers.

Among the BLS's biggest challenges over the past decade has been budget cuts leading to loss of purchasing power. During the federal budget sequestration years starting in 2013—the year Groshen was confirmed as commissioner—the BLS budget was cut in real dollar terms. While restored in nominal dollar terms by 2016, the BLS continued to lose purchasing power, cumulating to a $75 million loss (13 percent) since FY09, as shown in Figure 1. Abraham, Groshen, and Poole speak to the effects of these cuts in their responses below.

Coordinating and sharing data with the other two federal economic statistical agencies—the Bureau of Economic Analysis (BEA) and the US Census Bureau, both in the Department of Commerce—is another major challenge for the BLS. Being distinct agencies in different departments, these three agencies experience significant legal and bureaucratic hurdles to data sharing and coordination, hurdles that ultimately increase respondent burden and result in discrepancies between agency releases.

The current administration, as well as the prior two, have attempted to enhance data sharing for the three economic statistical agencies. In the early 2010s, the Obama administration launched a "data synchronization" effort, a revenue-neutral proposal to enable the BLS and US Census Bureau to better synchronize their business lists and broaden BEA's access to Internal Revenue Service (IRS) business tax data. Unfortunately, this initiative failed to gain momentum in Congress. The same is true of a Trump administration proposal—part of a larger government streamlining initiative—to

reorganize BLS, BEA, and the Census Bureau under the Department of Commerce "to increase cost-effectiveness and improve data quality while simultaneously reducing respondent burden on businesses and the public." The Biden administration, in their FY22 budget request for the Department of Treasury, proposes to provide BLS and BEA expanded access to some federal tax information. Congress so far has not acted upon this proposal in their FY22 appropriations bills.

While concerned about the many challenges the agency faces, the BLS experts also praised the current BLS leadership, the agency's nimbleness during the pandemic, and the progress BLS has made over the last few decades to better meet data-user needs.

The bottom line is that, despite having a high degree of autonomy and stature, the BLS is hampered in its efforts to track the rapid changes in the economy as well as it would like; meet data-user demands for more timely, granular data; and address declining response rates. Its future relevance may be at stake. Nonetheless, BLS continues to produce widely respected and monitored gold-standard data on employment, working conditions, productivity, and inflation.

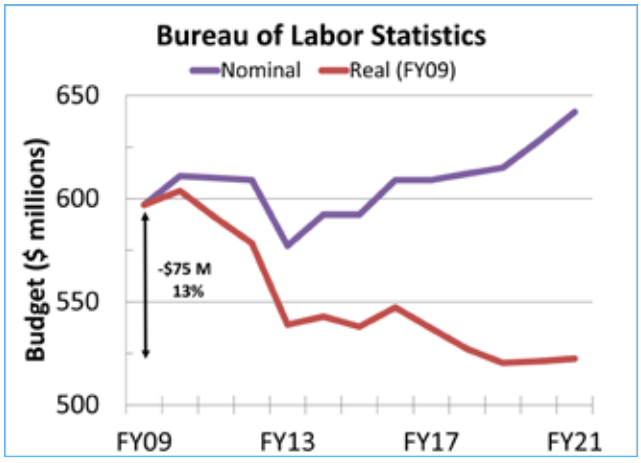


Figure 1. Bureau of Labor Statistics budget in real and nominal dollars

Abraham

Groshen

Poole

BLS is best known for its monthly employment reports. What are the bureau's next two highest-profile programs? What program do you wish were better known?

Katharine Abraham

Right now, lots of people are paying attention to inflation, so the Consumer Price Index is in the spotlight. Because of the ongoing discussion about how difficult employers are finding it to recruit people, the Job Openings and Labor Turnover Survey (JOLTS) data also are getting a lot of scrutiny.

Erica Groshen

I agree with those two and note that high-profile depends on current events and audience. Financial markets always look at the Consumer Price Index, while others' attention waxes and wanes. On BLS's website, the part with the highest usage is the *Occupational Outlook Handbook,* which provides career guidance on hundreds of occupations. That's a different audience than we normally think about for BLS, but it's a big one and speaks to whether we have a mismatch in skilled occupations. And we cannot forget BLS's productivity measures. While these releases don't generate many headlines, they are truly fundamental because they drive key conversations and decisions about policy and the well-being of the country.

Ken Poole

I'd like to turn to what BLS could be better known for and, specifically, the broad interest in localizing BLS's products on employment, productivity, inflation, and more. For unemployment statistics, prior to the pandemic, people weren't paying as much attention to the BLS data because the unemployment rate was low and wasn't changing much. The pandemic brought a dramatic change. Not only did people want more timely and readily available data, but they also wanted employment data that was more granularized to their communities—whether geographic or demographic.

This interest also applies to productivity and pricing data. I'm involved with groups trying to figure out the impact of automation on productivity. Automation also, of course, has ramifications for employment projections, which brings us back to the *Occupational Outlook Handbook*—a data source that relies on reliable projections.

More localized products from BLS would not only satisfy data-user needs, but they would also be a huge boon to economic development and job creation in communities across the US. I think it's really important for us to use the current opportunity to help promote innovation and reinforce the innovation that's been going on.

What are the biggest challenges facing BLS?

Ken Poole

Relevance. With demand for data way up and private sector providers in the marketplace providing more timely and responsive data, BLS is increasingly not being seen as the primary source. That's not to say the data available from those other sources is more accurate or valuable by any stretch of the imagination. But I think that's the biggest challenge—how do you balance the need for being the gold standard for data with the marketplace demands for immediately available data to drive decisions?

For example, ADP gets its employment data out before the BLS employment outlook to influence markets and give its customers advance warning about government reports. There's also the job openings and closing data that Burning Glass puts out. These private vendors are becoming direct competitors with BLS as a go-to source for data. While we can argue over the relative reliability of what these vendors produce, those entities are gaining visibility because they're getting data out there quickly. It's affecting the BLS brand.

People love JOLTS, for example, but the problem in my community is that it's experimental at a state level. We really need it at a labor market–area level. The BLS employment projections program is done at a national level; we're reliant on Employment & Training Administration (ETA) funding to support the production of substate-level information, and those projections are not done consistently across the states. In the current environment, policymakers want and expect real-time, localized data. They view BLS as the historian of data. Even the most current disaggregated data being released from BLS is three or six months old. Well, look at how quickly that pandemic hit us and how differently it affected various groups. These kinds of challenges are going to continue to affect how people perceive BLS and the value of BLS.

Katharine Abraham

This is a challenge not just for BLS but for all the statistical agencies. They're no longer the only entities collecting and disseminating data. But the existence of labor market data from other sources doesn't make the BLS data any less relevant. One of the first things ADP, Burning Glass, or any of the other private sources of labor market data will mention when they talk about their data is that they are confident in their numbers because they track the BLS numbers. There is definitely a role for the gold standard data.

That being said, all the statistical agencies need to be thinking about what their role is. In my view, it's not necessarily just producing the official statistics we all know and love. Yes, there's a concern that, if a statistical agency puts out products that are timelier or more disaggregated but based on incomplete data, it might undermine the brand. I am confident, though, that there's a way to differentiate between the official statistics that are the gold standard and more experimental measures that the agency has produced because they meet a need for more timely or more granular information.

Erica Groshen

I agree. We need to figure out the best relationship between the gold-standard series (where consistency is really important) and more timely and relevant data that addresses the particulars of current situations. This could involve a sunset provision for the data programs, so people know BLS will not be asking these questions forever. Having the capacity to ask the new questions quickly, with no expectation that you will ask them forever, is important. Related to that is the challenge of incorporating new data sources into the existing programs without undermining their quality or consistency, or having to change sourcing every week.

Some internal challenges, besides funding constraints, include attracting and retaining top staff in an age when data science is so important. In the long run, it should help the statistical agencies to have more inflow and outflow of people between the private sector and public sector. However, this is a challenge for the agencies in the short run because of the current high demand for data science skills outside of government. Unless the agencies can better align compensation with the external market, this will be a problem in the long run, as well.

Another challenge is achieving the economies of scale and scope in operations to get the biggest bang for everyone's tax dollars while preserving independence. Since the passage of legislation in 2015, the statistical agencies are being required to try to achieve such efficiencies (by sharing computer systems, human resources, travel, and other such functions) within parent agencies instead of with other statistical agencies. This path risks undermining the independence of the agency and may not generate savings. Many needs of the statistical agencies are quite different from those of other parts of their parent agencies. Having the parent agency leadership (with their policy and political focus) in charge of shared services puts the statistical agencies in a potentially compromised situation. Whereas if

you shared services among statistical agencies, you'd have more commonality of mission and roles and could avoid the potential for political interference.

What has been the cost of the 13 percent loss of purchasing power over the last 12 years to the US economy and labor?

Erica Groshen

Early in my term, because of budget cuts, I had to eliminate the mass layoff statistics, the green jobs initiative, and international labor comparisons. Part of the reason for choosing these programs was that they were underfunded, so they weren't what they should have been to begin with. But the needs that sparked their creation are still out there. With more funding, BLS could have conducted the programs the right way. In another example, JOLTS was started on a bit of a shoestring, so the sample size was not as big as it should have been. Even with the limited sample, the program has proved quite valuable.

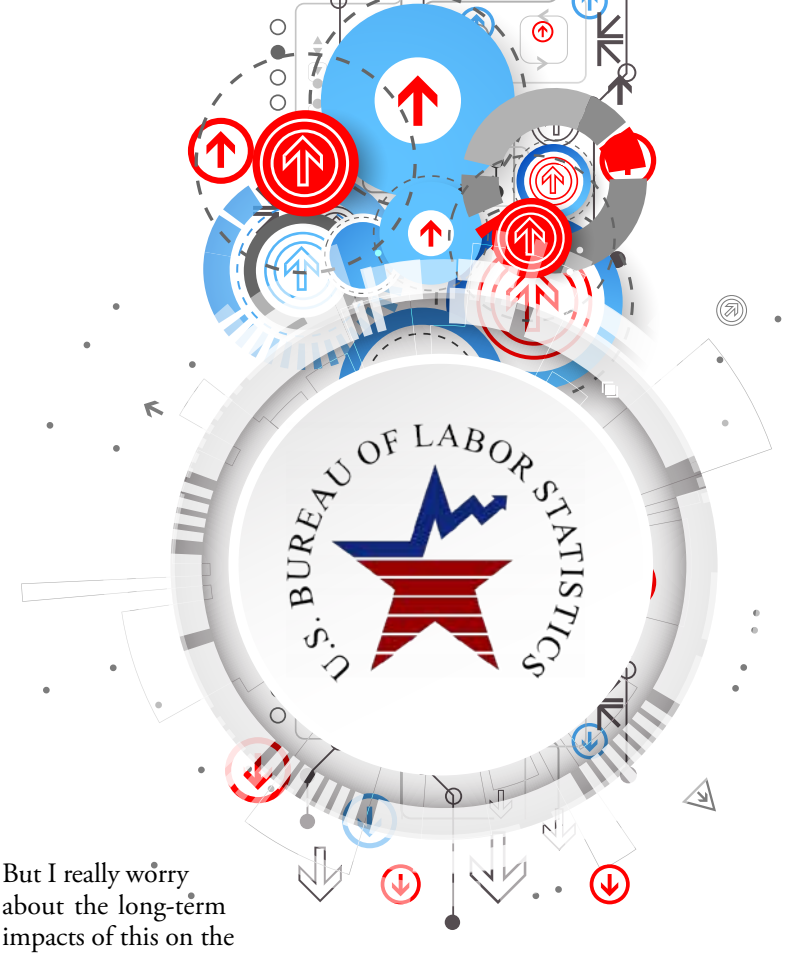
BLS has put forth initiatives to improve its timeliness and granularity, but these have not been funded. BLS has also been trying to modernize its Survey of Consumer Expenditures for two decades now—at a snail's pace because it has had to do so without additional funding. In one more example of "deferred maintenance," I think we all agree it's time to rethink how we do the occupational projections.

The overall point is that statistical programs need the bandwidth to do two things in addition to producing their regular releases. They must always be testing and validating improvements to meet data users' needs and keep pace with technological and economic changes. And they must have resources to provide resilience in case of potential disruptions such as weather events, strikes, pandemics, and so on. The loss of purchasing power has limited BLS's ability to innovate as much as they want to and narrowed their margins for resilience. So, the budget hit has been costly in terms of timeliness, frequency, granularity, agility, and relevance.

Katharine Abraham

A lot of what resource constraints do is lead to scrimping on the infrastructure—for example, not updating the computer systems, not giving your staff the appropriate training, or not funding the states at the level they need to do the work on the federal-state cooperative data programs.

Unfortunately, that can translate into negative effects on the quality of the data. The deterioration in data quality may not be very visible, and if the agency continues to produce the data products, users may not notice, especially not right away.



But I really worry about the long-term impacts of this on the quality of the data.

The increasing difficulty of collecting survey responses continues to be a big issue. I have been shocked by the magnitude of the increase in nonresponse in recent years. Even for the agencies' flagship surveys, it's getting harder and harder to collect survey data. The Current Population Survey response rate, for example, has fallen by roughly 10 percentage points in the last 10 years. We need to rethink how we're obtaining the information needed to produce the statistical data series users rely upon.

Ken Poole

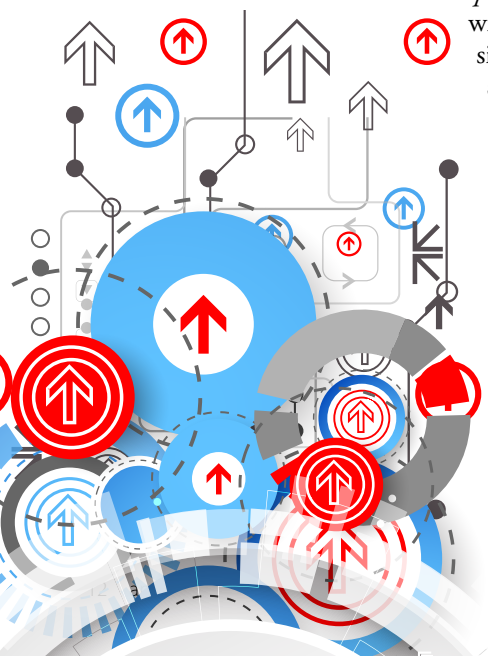
Over the past two decades, BLS has endured a period in which annual budgets were steady or received slight cuts. Year after year, we heralded these as small wins in an environment in which many other federal labor department programs were enduring massive budget cuts. However, we continued to operate so the public did not see the degradation in resources available. As a result, the impact of these decisions were not clear to the public. Sometimes you have to cut the services to get attention to the impact of these thousands of decisions over the years.

These cuts have been a particular concern to the state labor market information agencies. They have endured sizable budget cuts, not only from BLS but also from ETA—their two primary sources of income. So, some states are reallocating dollars from the general state fund to make up for the federal

funding shortfall, while other states simply don't have the resources to staff the data gathering and cleaning programs they once had. In some cases, the state labor market information agencies are basically hustling business from other state agencies to keep their staff in place because they don't have the resources, which may not be how we want to run our statistical agencies charged with providing public data. That means those states are focusing on doing whatever the immediate project is, as opposed to making sure we all have the high-quality, well-maintained data we need over a period of time.

BLS's occupational projections data is a great example of how ETA had to step forward to ensure state data remained available when BLS cut programs in the past. State and local projections must wait until BLS releases national projections before the state agencies can even begin their work. As a result, the data is often out of date by the time it comes out. In addition, due to cost constraints, survey data is limited so emerging occupations are lumped within existing occupations far longer than might otherwise be necessary. It also takes years until emerging occupations show up in our data.

The perfect example is data scientists. How long did it take us to get data scientists into the occupational system? Think about cyber networks and cyber security issues. Where do you find cyber information security specialists, and how can we bubble up these and other rapidly changing occupations more quickly? We need to invest in our understanding of the rapid changes to our occupational framework, and maybe this notion of skills taxonomy will help us see these changes more rapidly. The *Occupational Outlook Handbook* is the most widely read document on the BLS website and it relies on BLS occupations survey data, so it is the most significant resource BLS provides to the general public. That means occupation data guiding students toward new careers needs to reflect those emerging new opportunities.



Declining response rates and small sample sizes sound like technical issues data collectors must address, but they affect data relevance. We need to connect these issues in the minds of policymakers to the similar problems that face opinion polling that electoral candidates experienced in the last couple of elections. We can't assume we're getting to the people we're supposed to be getting to. That's why we worry about the quality of polls and why we should also worry about the quality of survey data about labor markets. If we can begin to connect that dot between this issue and the value of administrative data sharing as an alternative to surveys while also ensuring more transparency and better governance of data sharing across the statistical agencies, we might have a lever to help us overcome legal and policy barriers to the effective use of administrative records.

How would you broadly assess the state of BLS?

Erica Groshen

BLS continues to function well. Yet, serious impediments prevent the agency from doing many things to further improve its products. Besides funding, there are the very real threats to BLS autonomy I already mentioned. Further, their coordination with their sister economic statistical agencies, the US Census Bureau and BEA, is limited. Federal law prohibits the Census Bureau from sharing business information based on IRS data. So, the agencies' sampling frames are different and they get different results for the same concept, such as industry employment. They also have different approaches to aggregating industries and occupations and use different geographic areas. Such inconsistencies make it difficult to combine data produced by the three agencies. Lowering the barriers to coordination among these agencies would address this. It would also facilitate more reuse of data, which would ease the burden on data suppliers and the modernization efforts.

Katharine Abraham

Building on Erica's point, the best path forward may actually be to combine BLS with the BEA and the economic statistics part of the Census Bureau. In the past, I have resisted this idea. Any kind of reorganization in government is disruptive, and I did not believe the benefits would be worth the cost. But I've come around to thinking that, in this new world where we increasingly want the agencies to be making use of new types of data to do new things, it is unlikely to be efficient to have the agencies all pursuing those goals independently. Rather than the BLS, the BEA, and the Census Bureau each contacting the same potential sources of useful economic data, for example, a combined agency could approach those sources

It's always been hard to generate support for the BLS that translates into advocacy.

on behalf of the statistical system as a whole. You obviously would need legislation to create a new combined agency. Perhaps you could get to the same place with better coordination among the separate agencies, but I'm skeptical.

Erica Groshen

I agree. In order to do that, you need an empowered person or entity over all three agencies. It would make a lot of sense to have the chief statistician of the US take on this role. The chief statistician would oversee some key federal statistical agencies, rather than leading only a small staff in the Office of Management and Budget. Furthermore, you would want this combined agency to have clear independence, that is, not to report to a Cabinet secretary.

Ken Poole

I'm going to play devil's advocate. One of my biggest concerns in the way the agencies currently operate and may very well be exacerbated in a situation where the agencies are together is that they seem to be somewhat divorced from who their client base is. BLS has historically been primarily oriented toward OMB or Congress and toward the need for national data that contributes to the principal federal economic indicators, but one of the key reasons why statistical agencies are successful in getting resources and innovating is that they respond effectively to a broad customer base.

Often innovative data programs come from program offices, rather than the statistical agencies, because they are responding to immediate customer needs. For example, one of the biggest innovations in employment-related data provision is the Occupational Information Network (O*NET). It's operated out of ETA, rather than BLS, even though ETA still respects BLS as the primary source for methodology and methodological improvements. O*NET is driven by the ETA customer base—workforce agencies, education partners, or economic development agencies.

So, this tension exists between what the statisticians or technocrats in charge of the statistical agencies think innovation is needed and what their customer base desires. However, the customer base has Congress's ear. So, if we were to go to a model in which the federal economic statistics were aggregated into one place, opportunities for the customer's voice to be heard would need to be robust and transparent.

I think this disconnect is part of the struggle the statistical system has had over the last two decades, where the statistical system has been fighting over budget scraps when we had this push and pull of different policy priorities. So, I think there's something to be said here for how we engage the customer in this process in a way that helps the agencies become as responsive as possible within their budget constraints. The challenge is that these customer demands don't occur in many conversations about statistics. Instead, they occur in the natural ebb and flow within program agencies about policy priorities. This engagement with program agencies will need to balance the need for independence in the political environment that we're talking about with the need for maintaining relevance for the statistical agencies.

This disconnect is also the reason I would describe BLS as at risk. They don't have the depth of customer base many other programs have. Congress and the administration, for example, are deliberating a major investment in infrastructure and nobody seems to be buying the argument about data being a core infrastructure. I think that's because BLS just doesn't have the community-based advocacy network it needs to make a big difference on Capitol Hill.

Katharine Abraham

Speaking to the advocacy network, the Friends of BLS is a great innovation. It's always been hard to generate support for the BLS that translates into advocacy. When I was at the BLS, we used to talk about the support for BLS being a mile wide and an inch deep. There are a lot of people who appreciate BLS data, but the BLS is hardly anyone's top priority.

On a positive note, BLS benefits from good leadership. Commissioner Beach is focused on the right issues and doing a really good job in his efforts to move the agency forward.

Ken Poole

I agree with you. I've been monitoring BLS from the user-community perspective for more than two decades and I have seen an evolution in how I interact with BLS and how BLS interacts with the broader community. I think there have been really important shifts in the culture of BLS's leadership to be more responsive to economic and societal dynamics. BLS also seems to be shifting from being perceived as a survey agency to more of a data aggregation, collection, synthesis, and interpretation agency. I think this is a healthy move despite the political risks of such innovation. To reiterate my point about relevance, I don't think BLS can remain relevant unless it makes these kinds of changes.

If BLS were provided unlimited resources, what would you like to see the agency do first?

Katharine Abraham

The quality of the national data the BLS produces obviously is important, but I also would urge the BLS to expand its efforts to produce more granular data on a timelier basis. What is the saying? All politics is local? Several years ago, I served on a panel convened to evaluate the Census Bureau's program of annual economic surveys. We brought in a group of state and local officials to talk about the economic data produced by the Census Bureau. Like the BLS, the Census Bureau produces a fair amount of state-level data. But what these local officials really wanted was data for their metropolitan area or their county or, even better, even more disaggregated geographies that they could aggregate up to geographies that were meaningful to them. And they wanted data that was current, not data that refers to what happened a year or two ago in some cases. We are just so far from being able to meet this demand.

Erica Groshen

The national needs are important. We don't want to jeopardize the quality of the national statistics, but I think there's a huge demand for not just state, but local area, data. With all this private sector information—the ADP data, Burning Glass data—if BLS worked with some of those private sector vendors and figured out ways to produce modeled statistics that took that data as input, it would be possible to produce really valuable information that's not being produced now. It would be a huge investment required to do that.

And so let me take it one step further. I would give BLS ongoing access to enhanced unemployment insurance (UI) wage and claims records, with occupation, hours, location, and demographics added. These are the recommendations of the Workforce Information Advisory Council's (WIAC) 2018 report (*https://bit.ly/3D3zs8l*). Implementing this will allow BLS to do everything we're talking about. Such an effort would complement a business-led effort to create standardized employment records in companies. Funding the states to curate them appropriately and aggregating this data would be immensely powerful, yielding true economic statistics (such as improved mass layoff statistics) based on the wage and claims records at a more granular level than we have seen before. With modeling, one could also address the timeliness issue. And BLS would be able to eliminate or reduce the scope of at least two of its current surveys.

Ken Poole

I completely agree with Erica that BLS having access to enhanced UI wage record access would address many of the BLS challenges. We'd also be learning a lesson from the private sector, which models existing data to deal with privacy and confidentiality issues and create projections that give them a better sense of what's happened recently, what's happening now, or what's likely to happen in the not-too-distant future—all in a much more seamless way. Currently, the way the system is set up, it takes so much time to do that kind of data alignment, that it's not worth using public data. So, the user gives up on federal data and basically says that to actually use data in a way that is most relevant, they must go through a private intermediary to do so.

I also don't want to lose the idea of increased sharing of business lists between the Census Bureau and BLS already mentioned, because one of the most frustrating things for the user community is for one data source to tell us there are 300,000 manufacturers in the US and another data source to say 400,000. They're both reliable federal statistical agencies saying different things. And it's because they can't share their data to align their answers. It creates unnecessary confusion, and that shouldn't be the user's problem. We should figure out a way to solve that problem among the statistical agencies.

O*NET really needs to be upgraded to be structured and standardized. Operated out of the ETA, it's essentially a survey of jobs and the skills required to do those jobs. It isn't a BLS product, but it ought to be because it's one of the most powerful tools we use in this day and age and because it digs down beyond counting numbers of jobs and into the

makeup of those jobs. This is related to a BLS initiative underway to design skills taxonomy, because we're beginning to track and ask questions about skills more so than about industries or occupations in terms of worker preparation.

In short, BLS needs to develop data on skills, and I'll reiterate the need for localized, real-time, high-frequency, forward-looking labor market information.

The idea of BLS making more use of private sector data has been a strong theme in this discussion. How feasible is this?

Erica Groshen

There are a lot of issues along that path. Will those organizations charge the statistical agencies for the data? Will the data continue to be available? Will quality be maintained? Will companies use pre-release information to front-run statistical releases? Who is responsible for privacy? These are resolvable, but it will be easier to do so if the agencies can approach the private entities as a united front.

Ken Poole

Several other agencies are out there buying this data, independent of the statistical agencies, without understanding how it might relate to the public statistics. You don't have the expert opinion, or you're paying for additional expert opinion, to figure out how their stuff is applicable in the various circumstances. There is something to be said for the statistical agencies using it to help model their data. Even if it does cost money, it would be better for BLS to buy it once than to pay for it 1,000 times through the various other federally funded programs.

Readers of this discussion may infer the private sector may one day supplant the need for some of BLS's programs. Is this realistic?

Katharine Abraham

At least with respect to the core BLS data products, that's not realistic. There will continue to be a need for publicly available economic statistics that everyone agrees are gold standard measurements. Even with regard to more detailed and more real-time data, statistical agencies like the BLS are likely to have an edge when it comes to producing statistics that are broadly representative. While information from private sources certainly can supplement the official statistics produced by the BLS and other statistical agencies, I don't see that information ever supplanting the official statistics. ■

2022 Research Fellowships at BLS, BEA Available

The American Statistical Association, in cooperation with the Bureau of Labor Statistics (BLS) and Bureau of Economic Analysis (BEA) under a grant from the National Science Foundation (NSF), will offer a senior research fellow program in 2022.

The fellowship at BLS allows research fellows to use BLS data and facilities and interact with BLS staff. More information is available on the BLS website at *https://bit.ly/36NzdRL.* or in the downloadable brochure at *https://bit.ly/3o7BzBQ.*

Application Deadline: **January 3, 2022**

The fellowship at BEA offers an opportunity to perform research there. BEA produces key economic statistics that influence government policy, forecasting, and business investment. Fellows will have access to BEA data and the expertise of BEA staff. More information is available at *https://bit.ly/3C3MSQm* or in the downloadable brochure at *https://bit.ly/3kGKOZ1.*

Application Deadline: **January 3, 2022**

Eligibility

An academically recognized research record and considerable expertise in the area of proposed research is required. US government employees are not eligible to apply. Applicants must be affiliated with a US institution.

Condition of Appointment/Benefits

Research will be conducted at the government agency. The stipend received is commensurate with qualifications and experience, and the term of the appointment is flexible. Fringe benefits and travel allowances are negotiable.

ASA Members Honor Three with Special Awards in Statistics

The 2021 International Science and Engineering Fair (ISEF) was held virtually April 29 – May 21 and, as they have for more than 30 years, ASA members volunteered as judges for the projects.

ISEF is the world's largest scientific competition, with millions of high-school students around the world competing to secure one of about 1,450 coveted spots. These spots are awarded to students who have advanced from local and regional fairs based on the quality of their projects.

Involvement in ISEF has been an outreach project of the ASA Council of Chapters since 1987. The council appoints an ASA-ISEF liaison, and the ASA provides a modest budget that includes travel expenses for one chapter member to travel to the fair.

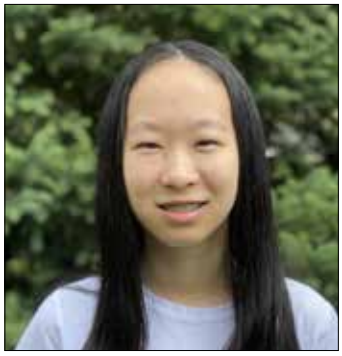
Judging for ISEF ASA Special Awards in Statistics traditionally has three rounds. In Round 1 for ISEF 2021 (held virtually), the statistics judges screened 1,430 projects by 1,842 students from 64 countries. These projects represented 21 scientific and engineering categories. Of these, 150 projects that incorporated sufficient statistical details were selected for more detailed review using a quantitative rubric in Round 2. Fifteen projects were selected for Round 3 interviews based on the quantitative Round 2 scores.

The three projects selected from Round 3 based on the best use of statistics using a quantitative rubric are:

Aditya Kendre

First Place ($1,500) went to Aditya Kendre, Cumberland Valley High School, Mechanicsburg, Pennsylvania, for the biomedical engineering project ENBM035: "Employing Adversarial Machine Learning and Computer Audition for Smartphone-Based Real-Time Arrhythmia Classification in Heart Sounds."

This project used a generative adversarial network (GAN) and transGAN subsystems to detect abnormal heart sounds to classify arrhythmia. Kendre also developed an app to be integrated into a smartphone. The project model testing, evaluation, and interpretation incorporated many key statistical approaches in an innovative engineering project.



Hannah Guan

Second Place ($1,000) went to Hannah Guan, BASIS San Antonio Shavano Campus, San Antonio, Texas, for the computational biology project CBIO008: "The Genetics of Human Aging: Predicting Age and Age-Related Diseases by Deep Mining High Dimensional Biomarker Data."

In this project, the correlation pre-filtered neural network model was developed and tested yielding the best prediction accuracy among all tested methods, including neural networks with LASSO regularization, elastic net regularization, and dropout neural networks.

Third Place ($500) went to Sarah Cao, Palos Verdes Peninsula High School, Rolling Hills Estates, California, for the Plant Science project PLNT016: "Analysis and Prediction of the Spread of Invasive Plant Species in Relation to the Changing Environment."

Sarah Cao

Virtual Complications

There were several traditional special outreach tasks that are easy for the judges to complete in person but required special handling to accomplish in the virtual fair. For in-person fairs, *CHANCE* and *Significance* magazines are distributed to students and their teachers who attend the statistics symposium; however, in 2021, no symposium was held.

Another rewarding task at in-person fairs has been the distribution of donated statistics books to winners of the top projects. A subsequent article in *Amstat News* will describe the solutions for continuing the tradition of distributing books during ISEF 2021.

In this project, the random forest machine learning model was used to determine the relation between the spread of invasive plant species and various environment factors and to predict future plant spread with promising accuracy.

In addition to the cash prizes, each winner received a one-year student membership to the ASA, including one-year subscriptions to *Significance* and *CHANCE* magazines.

Ten projects from students living in countries such as the US, Vietnam, India, and Russia received honorable mention.

Xinping Cui, president of the ASA Orange County/Long Beach Chapter, announced the three winners of the ASA Special Awards in Statistics during the ISEF 2021 Special Awards Ceremony, which was also virtual (*www.youtube.com/watch?v=2KDvEdehEMs* - ASA awards begin at 23:00).

One of the social features of the in-person fairs that could not happen with this year's virtual fair was the dinner traditionally hosted by the ASA; however, the ASA did have several "happy hour" Zoom sessions. The highlight of one of the sessions was special guests Susan Ellenberg and Ananda Jayawardhana, who provided their perspectives on early ASA involvement in ISEF.

For more information about ISEF, visit *www.intel.com/content/www/us/en/education/intel-education.html.* ■

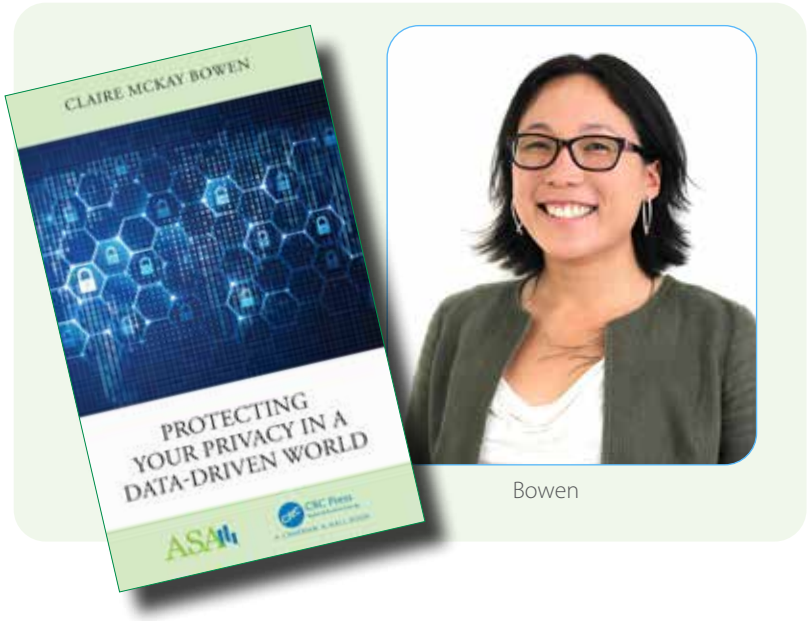
CALL FOR PAPERS

The Molecular Diversity Preservation International open-access journal ***Stats*** will feature a special issue, "Small Area Estimation: Theories, Methods, and Applications." The deadline for manuscript submissions is December 31.

Stats guest editor, Balgobin Nandram, is looking for papers in small area estimation that have strong theories, methods, and applications or a mix of these broad areas.

For information and to submit a paper, visit *www.mdpi.com/journal/stats/special_issues/small_area#info.*

Claire McKay Bowen Talks Data Privacy



Bowen

What initially drew you to write about the tradeoff of data privacy and the use of data for the public good?

There are two parts to this question. Why write a book on data privacy and why on the topic for the public good?

To answer the first question, there are *still* few communication materials about data privacy, or methods of safely releasing confidential data publicly while preserving the privacy of those who are in the data.

Years ago, as a naïve first-year doctoral student, I excitedly started my literature review on data privacy with a focus on differential privacy, a mathematically rigorous definition for privacy. But my excitement quickly transformed into frustration. I found only a few technical papers (especially within statistics) and even fewer introductory materials. I remember scouring the internet for *anything*. To make matters worse, some of the blogs, articles, and academic papers I could find had incorrect definitions, unclear explanations, or improper applications of differential privacy. What I found most useful were statistics and computer science professors' public PowerPoint slides or random white papers from federal agencies.

As differential privacy and other data privacy techniques gained in popularity, the privacy community created more technical papers, blogs, videos, and general privacy outreach. Fortunately, this increase in communication materials will result in fewer students stalking professors' web pages and generate a more accurate understanding of data privacy. Unfortunately, most of these materials still struggle to explain differential privacy and other data privacy concepts to a nontechnical audience. This became my initial motivation to write the book. We need more accessible written materials.

Why the focus on the public good? Although this version of data privacy may not seem as exciting as others, such as cybersecurity, it affects every person's life through major public policy decisions in the United States. When I wrote the book, the world was (and still is) experiencing a global pandemic that has caused severe economic and health public policy issues in most countries, including the United States. If researchers and public policymakers had access to tax and health data, they could better target and coordinate stimulus relief programs to help all American residents.

But many public policymakers do not understand the tradeoff between data privacy and public good. Therefore, I decided the intended audience for my book includes anyone interested in learning more about this area of data privacy without a mathematics background. Specifically, public data users, people working within the state and federal government who are not as

Claire McKay Bowen is the lead data scientist for privacy and data security at the Urban Institute. Her research focuses on developing and assessing the quality of differentially private data synthesis methods and science communication. She holds a BS in mathematics and physics from Idaho State University and an MS and PhD in statistics from the University of Notre Dame.

After completing her PhD, Bowen worked at Los Alamos National Laboratory, where she investigated cosmic ray effects on supercomputers.

In 2021, the Committee of Presidents of Statistical Societies identified her as an emerging leader in statistics for her "contributions to the development and broad dissemination of statistics and data science methods and concepts, particularly in the emerging field of data privacy, and for leadership of technical initiatives, professional development activities, and educational programs."

familiar with data privacy preserving methods, and public policymakers who want and need to learn more about data privacy methods to make more informed policy decisions.

Particularly in the context of the COVID-19 pandemic, how would you respond to the claim that a public health crisis trumps data privacy concerns?

Before even deciding if we want to release personal information for the public good, we should be aware of how this tradeoff is not distributed equally in society. In other words, we should determine if the decision we make, whether to use or withhold the data, doesn't create both worse privacy and health outcomes for certain individuals.

Specifically, underrepresented groups tend to experience higher privacy insecurity and are at a higher data privacy risk. These individuals, racial minorities and socioeconomically disadvantaged people, are also more likely to suffer from health policy outcomes.

We saw this during the pandemic in the United States. For example, in May 2020, the Navajo Nation surpassed New York City for the highest per capita coronavirus infection rate. When the nation started to lock down, most health advisories stated we should wash our hands for 20 seconds. However, the Navajo Nation could not comply with this health advice because more than a third of the population lacks running water. The Native American populations in my home state of New Mexico also suffered similar issues. The local pueblos didn't open to the public until early 2021, when the local government distributed vaccines.

To make matters worse, researchers predict people who tested positive for COVID will be at a higher data privacy risk. This means the inequity of personal privacy will widen even more for underrepresented groups, who have a higher rate of infection. So, when we have these conversations about whether we should sacrifice our personal privacy for the public good, we must acknowledge how these decisions may affect some more than others in our community.

With data sets getting bigger and more complex, what are the biggest privacy challenges we face going forward?

The greatest challenge as data sets get bigger and more complex is the lack of data practitioners who are both knowledgeable about data privacy methods and have the skills to apply those methods. This answer might be what you expected, but even for some "simple" data sets, a data user must identify which algorithms to use and where to insert them into the statistical pipeline.

Also, production-ready data privacy software is still limited, although more is becoming available. Many programs only implement a few methods, apply to certain types of data, and/or do not scale up well for larger data sets.

Simply put, we must overcome the human and computational resource limitations to tackle the other privacy challenges.

To learn more about Bowen's book, *Protecting Your Privacy in a Data-Driven World*, visit *https://bit.ly/3H9tZ2c.* ■

MORE ONLINE
To learn more about Bowen's book, *Protecting Your Privacy in a Data-Driven World*, visit *https://bit.ly/3H9tZ2c.*

MY ASA STORY

Jeri Mulrow | Applied Statistician

Mulrow

Hongyuan Cao, Jeri, and Renee Moore gather during the 2010 People to People trip to China.

The best thing about the ASA is it continues to grow and evolve as we, its members, adapt to meet the changing statistical needs of our times.

I *love* being a member of the ASA. I've had a wonderful career in statistics, and the ASA has been my professional organization all throughout, promoting the profession and practice of statistics. I joined the ASA when I was in graduate school at the behest of my professors at Colorado State. Since then, I've had the opportunity to serve in a variety of roles at both the local and national levels. All have been great experiences in their own ways and have offered me chances to grow my leadership and networking skills, meet amazing leaders in the field, and make many friends.

I was asked to share a specific story about an ASA experience I found the most rewarding. Thinking back over my many experiences, there are two that stand out to me that I think are worth sharing. Both occurred during my terms on the ASA Board of Directors, first as a Council of Sections Governing Board member from 2009 to 2011 and then as vice president from 2014 to 2016. It was deeply rewarding to be part of the governing body of the ASA, helping to set strategy, supporting presidential initiatives, reviewing financials, and learning more about all the programs and activities of the ASA.

One rewarding experience stems from the ASA as an international organization that supports statisticians across the globe. The ASA has an international representative to the board, plus programs that focus on international members and activities such as Statistics without Borders and the Committee on Scientific Freedom and Human Rights. I was fortunate to participate in the ASA People to People Ambassador Program to China in 2010.

Our statistical delegation consisted of 34 ASA members and 12 guests. We met with statisticians in China at five universities

and visited the National Bureau of Statistics of China. My quote in the February 2011 issue of *Amstat News* sums it up: "The ASA People to People Statistical Delegation was a truly fabulous experience. It was the perfect mix of professional and cultural events. … It was exciting to meet our Chinese counterparts and to learn more about the field of statistics in China. However, it was even more rewarding to meet and to get to know so many terrific ASA statisticians and their guests who were part of the delegation."

The other rewarding experience arose because of the ASA's involvement in K–12 education via classroom resources, publications in statistics education, guidelines and reports, workshops and webinars for teachers, and student competitions. I had the opportunity to visit the 2015 Meeting Within a Meeting (MWM) at JSM in Seattle. The MWM program is designed to enhance educators' understanding of statistics and provide them with hands-on activities they can use in their own classrooms to strengthen the teaching of statistics. I observed both middle-school and high-school sessions for math and science teachers who were excited to be at JSM and learn better ways to teach statistics. It was heartening to see their enthusiasm for teaching statistics and their appreciation of the ASA for support and resources. The ASA's involvement in educational activities lay the groundwork for a robust future of the science of statistics, and I am proud to support these activities as a member.

The best thing about the ASA is it continues to grow and evolve as we, its members, adapt to meet the changing statistical needs of our times.

Jeri Mulrow is vice president and director of statistics and evaluation sciences at Westat. ■

MY ASA STORY

Allison Grove
Research Consultant

. . . Learning from the Experiences of Others

Grove

I'm a new member of the ASA, having joined in 2019. With an MS in animal science, I've worked with animal science research for 25+ years, primarily at universities and branch research stations. In 2008, I started AG Research LLC, providing research support on a freelance basis to academic, industry, and nonprofit clients. I live in rural Montana and remote work is sometimes the only option for continuing a research career. While I enjoy all aspects of a research project, I find the statistical analysis the most rewarding. A couple of years ago, I started taking online courses toward earning a graduate certificate in applied statistics.

Working remotely means I needed to find a way to keep in contact with other researchers and statisticians. Over the years, I had taken some workshops where Nora Bello was an instructor and reached out to her with questions. She was kind enough to reply and suggested I join the ASA. I like the way sections are organized by area of interest and joined the Consulting Section. When the COVID-19 pandemic hit, Isabella Ghement started a biweekly Zoom meeting for consultants to support and encourage each other. The topics of the discussions vary and have been a great way to stay connected and learn from the experiences of others.

There is another way I have been encouraged by being a member of the ASA and that is through a column in *Amstat News* (June 2019, Issue #504). In this issue, author Jillian Payne described how hiring managers desired candidates who have a background in art or music in addition to being skilled data scientists. This creativity helps with problem-solving. I am also a fiber artist and enjoy weaving but had been afraid to admit this to some of my science colleagues. After reading this article, I realized I didn't have to be and have enjoyed the "Pastimes of Statisticians" articles. I have also discovered many other fiber artists with scientific training.

Allison Grove is a statistician and scientific writer at AG Research LLC, Freelance. ■

THE JEDI CORNER

Caucus for Women in Statistics Celebrates 50 Years

Donna Brogan, Nairanjana Dasgupta, Amanda Golbeck, Wendy Lou, and Motomi Mori

The JEDI Corner is a regular component of Amstat News *in which statisticians write about and educate our community about JEDI-related matters. If you have an idea or article for the column, email JEDI Outreach Group member Cathy Furlong at* communicate@datascijedi.org.

The Caucus for Women in Statistics (CWS) is celebrating its 50th anniversary this year. Formed in 1971 before "JEDI" became an acronym known to all, CWS has been on the forefront of giving a voice to women in statistics and offering support, mentorship, friendship, and a safe space for discussion about challenging and often difficult topics faced by women.

The Past

It all started with Donna Brogan and a group of courageous women realizing there was a need for women in statistics to have a voice. The following excerpt is Brogan's story about how it started, which she shared parts of during the JSM 2021 session, "Celebrating the 50-Year History of the Caucus for Women in Statistics: Origins, Accomplishments, and Challenges."

> In 1967, I received my PhD in statistics from Iowa State University and accepted an offer of assistant professor of biostatistics at The University of North Carolina School of Public Health. I found myself on a steep learning curve about living and working in the professional middle class, a stark contrast to my working-class family background.
>
> In addition, I joined a women's liberation consciousness-raising group in the late 1960s in Chapel Hill; these women profoundly influenced my life. I finally realized that accumulated prior episodes of sex discrimination in my life were not personal or unique to me but part of society's structure, including rigid sex roles for women/girls and men/boys. Several group members had formed a women's caucus in their own career fields. I followed their example and decided that a women's caucus was needed in statistics, a male-dominated discipline at the time.
>
> At the Joint Statistical Meetings (JSM) in New York City in summer 1969, I attached flyers to the inside stall doors of women's bathrooms in the meeting hotel, asking women who were interested in discussing a possible women's caucus to meet in my hotel room. Seven to 10 women attended over three evenings and agreed that a caucus was needed. Afterwards, Dr. Anita Bahn and I discussed with John W. Lehman, then executive director of the American Statistical Association [ASA], our interest in forming a women's caucus to be affiliated with ASA; he was very supportive.
>
> At the December 1970 JSM in Detroit, joint with societies in economics, about 40–50 persons attended a meeting to discuss details of a caucus vision, mission, and structure. We decided to focus the planned caucus on statistics and recommended that the American Economic Association form its own women's caucus. In May 1971, the ASA Board of Directors recognized the Women's Caucus as an affiliated nonprofit organization, including participation in JSM (e.g., an invited paper session and space for a business meeting).
>
> The Women's Caucus made its official debut at the August 1971 JSM in Fort Collins, Colorado. Its sponsored invited paper session focused on equal opportunity for women in employment, with a lively question and answer session among 150 attendees. … I served as caucus president for three calendar years, 1971 thru 1973. … Our initial name was Women's Caucus. I don't remember when the name was changed to Caucus for Women in Statistics (CWS), but I like the current name better."

When CWS was founded in 1971, Donna and her cohort naïvely thought its life would be time bound. They assumed gender discrimination in statistics, employment, career development, mentoring, promotion, and recognition would soon be gone. Fifty years later, we take stock and realize that, although much progress has been made, discrimination still happens under different shapes. Hence, CWS has emerged as an advocacy group for women and other underrepresented groups and is vocal in the Justice, Equity, Diversity, and Inclusion (JEDI) drive.

The Present and Future

CWS is a self-reliant advocacy group and works across several organizations without being part of any. It partners with the ASA and often works closely with the Committee on Women in Statistics (COWIS) within the ASA. CWS also has ties with other statistical organizations such as the Institute of Mathematical Statistics, Statistical Society of Canada, International Statistical Institute, International Biometric Society, and International Chinese Statistical Association. It has played important roles in establishing the Gertrude M. Cox Scholarship, COPSS Florence Nightingale David Award, JSM First-Time Attendee Orientation, and Women in Statistics and Data Science Conference (WSDS). It also works every year to help increase the proportion of ASA Fellows who are women.

In the last few years, several awards have been established, including travel awards to help women attend conferences. Wendy Lou, 2020 CWS president, created the CWS Societal Impact Award to recognize statisticians whose contributions have influenced communities at large in terms of advancing social justice, diversity, equity, and inclusion through their research, teaching, or service.

Adding to the celebration of the CWS 50th anniversary this year, the first CWS Societal Impact Award winner was announced during the CWS business meeting at JSM 2021. Melody S. Goodman, a professor at New York University, was the inaugural recipient of the award. Her citation reads as follows: "For her outstanding dedication to improving health disparities through education in public health and data literacy."

In an effort to remember and honor the past while staying current in their mission, Tomi Mori, 2021 CWS president, initiated monthly Madam Presidents Happy Hour (*https://bit.ly/3CWo1zj*) hosted by past presidents. It is open to all and allows those who participate to get to know past CWS presidents and connect with others. Further, the CWS Past Presidents Interview Series is on YouTube at *https://bit.ly/2Yz3UZ6.*

Further Reading

Golbeck, AL. 2020a. Presidents of the Caucus for Women in Statistics, 1971–2021. *Significance Magazine*. *https://bit.ly/3bUdatZ*

Golbeck, AL. 2020b. Supporting an Inclusive Community: A Caucus for Women in Statistics. *Significance Magazine*, April 2020:42-44.

Pierson, Steve. Undergraduate Statistics Degrees Up 474% Since 2020. *Amstat News.* October 2021:10–16. *https://bit.ly/3BYMxP6*

Langer Research Associates. American Statistical Association 2020 Work and Salary Survey. *https://bit.ly/3F5atSR*

CWS has provided great role models for women in statistics. About three-quarters of prior caucus presidents are ASA Fellows. The caucus leadership is becoming more diverse in terms of racial and ethnic backgrounds, mirroring the growing diversity within our profession.

There is still more to do to advance women in statistics. The October 2021 issue of *Amstat News* shows that, among recipients of doctoral degrees in statistics and biostatistics, the percentage who are women was slightly below 40 percent in 2020 while women make up less than 20 (*https://bit.ly/3qkMVoT*) percent of doctorate earners in the computer sciences, physics, and mechanical engineering. The ASA 2020 Work and Salary Survey (*https://bit.ly/3F5atSR)* shows that more men than women are full professors (43 percent vs. 25 percent) and tenured (55 percent vs. 32 percent). And women in all sectors tend to make less, both in base salary and total employment income.

Join Us

CWS provides a forum for sharing stories. As Arlene Ash, 1986 CWS president, stated, "The Caucus for Women in Statistics is a safe space to speak out and be women, but not be entirely separate from the rest of the professional community." She also said, "The Caucus for Women in Statistics must continue to help us address questions of particular interest to women while welcoming their partners and all fellow travelers who support our goals."

The Caucus is a 501(c)(3) organization (following the US Internal Revenue Code), a nonprofit organization established for charitable, educational, and/or scientific purposes. All who are interested in the CWS mission are welcome to join. For more information, visit the CWS website at *https://cwstat.org* or contact Jessica Kohlschmidt, CWS executive director, at *jk.kohlschmidt@gmail.com.* ■

@WomeninStat + Rotating Curators = Success

The ASA Committee on Women in Statistics and Data Science has a rotating curator every month for its Twitter account. We wanted to know how the committee manages its rotation and what engagement strategies its members use to make @WomenInStat so successful, so we sat down with **Lucy D'Agostino McGowan**, chair of the Committee on Women in Statistics, and **Hannah Mendoza,** @WomeninStat rotating curator coordinator, and asked the following questions:

Hannah Mendoza, you handle the behind-the-scenes Twitter account for @WomenInStat. How did that start?

Hannah: During my final year of undergrad at Wake Forest University, I was a member of the D'Agostino McGowan Data Science Lab, where Lucy (my professor) advised me on my senior thesis project. I studied mathematics, but I wanted to do a project on the statistics side, as I had taken a few courses in statistics that piqued my interest. Lucy's focus on statistical communication resonated with me. During a lab meeting one evening in Fall 2020, Lucy asked if anyone would be interested in helping out with organizing and facilitating transitions between rotating curators for @WomenInStat. I volunteered because it seemed like a great way to learn from other women and minority-gendered people in the field through their diverse backgrounds and interests, as well as from how they communicate effectively and engage others on Twitter.

What prompted the committee to join Twitter?

Lucy: Last year, when Stephanie Hicks was the chair of the Committee on Women in Statistics, we were looking for ways to engage the broader community more. Inspired by the R-Ladies rotating curator account (@WeAreRLadies), we decided to kick off one of our own! It has been an amazing team effort. Hannah handles a lot of the behind-the-scenes work along with Stephanie and our current chair-elect, Eunice Kim. And, of course, we rely on all the amazing curators who have made it such a success!

@WomenInStat has more than 19,000 followers. What do you attribute the success to?

Lucy: I think it is really a testament to the behind-the-scenes team and all the amazing people we've had come through as curators! Since our kickoff in June of 2020, we've had more than 60 curators ranging across academia, industry, and government from varying backgrounds and at all levels of experience.

Hannah: The community—curators and followers—is super encouraging and passionate, and I think that makes @WomenInStat just a rather pleasant and inspiring place to be on the internet! And, of course, getting to host such a wonderful and diverse set of curators has drawn an audience that wants to learn about others and what excites and matters to them.

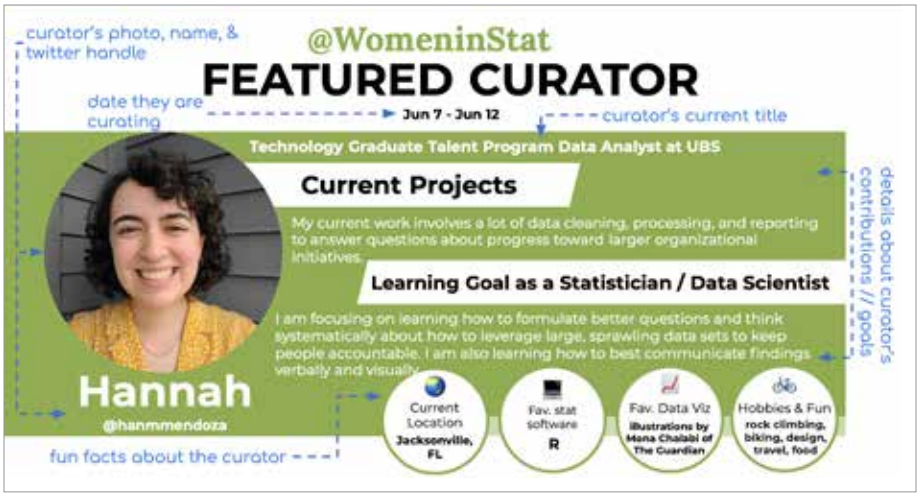
Describe your monthly plan. What actions and tools do you use to organize and manage the curators?

Hannah: We keep it simple and use GitHub and lots of tools in the G Suite! GitHub helps us centralize information for admins on managing the account and link to our resources such as a curator guide and our schedule (in Google Sheets). Curators sign up via a Google Form, so we have the survey responses in Google Sheets. We use these responses to generate an introductory graphic for each rotating curator (in Google Slides), which we tweet out from the @WomenInStat account at the beginning of each week to announce that week's curator (See Figure 1). We also monitor these responses for new sign ups so we can add them to our schedule. We communicate with curators through a central email (Gmail) to facilitate scheduling and sharing instructions.

Another super important tool for us is TweetDeck, a platform for managing Twitter accounts with multiple contributors or admins without having to share login information between accounts. This makes the weekly hand-off between curators really simple!

You can also schedule future tweets through TweetDeck, which is useful for us as admins and for our curators. One way we use this feature is to facilitate new curator sign ups. Every two

weeks, we have a scheduled tweet as a reminder of the purpose of @WomenInStat: to highlight and uplift voices of women in statistics and data science! And we share a link to the signup form to encourage others to curate for a week. This link is also in our bio for people to learn more and sign up! Our curators are also really great at encouraging others to sign up, no matter their background or experience level, which has helped create a culture around supporting and uplifting each other.



Example introductory graphic for rotating curators

What tips can you offer sections, chapters, or individuals who would like to become active on Twitter?

Lucy: I am a huge fan of these rotating accounts, both from the perspective of the groups that run them and as a follower of several! They offer a great way to engage with the community and figure out what others are working on. I hope more sections and chapters think about using this model in the future. For individuals, I have found a great way to get started is to try creating short summaries of recent happenings in your field. I started out by tweeting about talks at conferences I was attending—it was a great way to practice distilling complex information down into a few short tweets. I've always found that I understand topics better when I try to teach them—tweeting even more so since it requires pulling out the big ideas and trying to squeeze them into 280 characters (and finding the perfect emojis // gifs to get the points across!).

Share with us your favorite go-to websites for learning about social media.

Lucy: We have some tips for curating that we adopted from the R-Ladies documents in our GitHub repository (*https://bit.ly/3EVjQEk*). Shannon Pileggi gave a poster at the Women in Statistics and Data Science conference hosted this past October that gave some great tips for curating (*https://bit.ly/3FgNnc7*). In 2020, Jessica Lavery moderated a Caucus for Women in Statistics Lunch & Learn on creating a professional online presence and navigating social media. She has a great write-up about it on her blog (*www.jessicalavery.com/blog_cws_social_media_webinar.html*).

What topics get your followers 'talking'?

Lucy: Daniela Witten curated for us last year and tweeted about being a parent in academia, academic privilege, and training the next generation of statisticians. She also had an awesome thread, "It's just a linear model!", where she walked through a bunch of models and explained how just about everything boils down to a linear model. These topics were all super popular and generated a ton of great discussions.

I took a peek at the data from the last few months to see what generated the most chatter. We had a great thread on math and democracy (*https://bit.ly/30fthQj*), "Cheat sheets" for coding in R (*https://bit.ly/2YuwRoT*), and one with resources for learning causal inference (*https://bit.ly/3wsRCOs*). In general, it seems when the curators tweet about things they are excited and passionate about, engagement follows.

What do you consider the biggest benefit of being on Twitter?

Lucy: Community is so important—it makes life worth living! Getting to hear from other people in the field about what they are excited about these days is such a joy, and Twitter is one tool to facilitate this. It is definitely not perfect, but I have really loved meeting and getting to know people I otherwise wouldn't have had the opportunity to hear from.

Hannah: And so many communities already exist on Twitter! So, using Twitter lets us tap into and enrich existing communities through hashtags (#rstats, etc.) and when people share our curators' content to their own audiences and add their own insights. Twitter's infrastructure is great at facilitating learning and enabling discussion between people who may not cross paths otherwise. Using Twitter helps to make these communities and opportunities to learn and share resources more accessible. ■

MORE ONLINE
Follow @WomeninStat on Twitter and learn more about curating. Sign up to curate at *bit.ly/wsds-rocur.*

COMMITTEE SPOTLIGHT

Statistics and Disability (CSD)

Anja Zgodic, Sarah Zlatkovic, and Chuck Coleman

Valor, Erin Chapman's service dog, meets committee members during a recent Zoom meeting.

In your own words, what is the purpose of your committee?

Anja: The Committee on Statistics and Disability (CSD) was recently reactivated after a break. Based on conversations between a diverse group of members and allies, we have formulated an updated charge that reflects the values and direction the CSD aims to take.

First, the committee's goal is to serve the disability community within the ASA by working to improve accessibility and remove barriers to participation. Second, we aim to provide guidance to statisticians and the broader community on best practices and universal design to improve accessibility to statistics for the public and future statisticians. Finally, the last element of our renewed committee charge is to support and empower ethical statistical research on disability and accessible statistics education for people with disabilities, both within and outside of the ASA.

In my own words, I would say our new charge aligns with our strong focus on promoting inclusivity, accessibility, diversity, and their intersection.

Sarah: We are primarily concerned with improving access for those with disabilities. This means helping to facilitate participation in ASA functions and activities and providing guidance on making statistics accessible to the general public and future statisticians. Statistics is for everyone.

Why did you accept the position to chair and co-chair the committee?

Anja: I joined the CSD for the opportunity to actively participate in initiatives aimed at reducing barriers experienced by various groups of individuals. Whether it is within industry, academia, or the public realm, conversations about diversity and inclusivity have been increasingly taking place. These topics have rightfully been emphasized as important and I felt like I not only wanted to educate myself on systemic barriers, but also concretely engage in some form of advocacy.

Sarah: We recently re-energized this committee, and I was excited to be offered the opportunity to help. Advocating for individuals with disabilities is something I'm passionate about—it permeates my research and my professional and personal life.

Chuck: Since I received my autism diagnosis in 2007, I have been an active autistic self-advocate. The CSD is now one of my advocacy channels. From interactions with fellow Census workers, I have become aware of the needs of other disability communities. One common thread exists between the various historically disadvantaged communities: the need for diversity and inclusion and so on and the many barriers to overcome. The benefit for all lies in the ability to draw on the abilities and experiences of all of us and in the concurrent reduction of resentment and hatred.

How often does the committee meet to plan activities?

Sarah: We meet monthly via Zoom and stay connected through Slack. We utilize assistive technology like Otter.AI captioning during meetings to help make it easy for everyone to participate.

Anja: In addition to our monthly virtual meetings, CSD members frequently engage with each other

on platforms such as Slack. We often share articles, resources, and other items of interest, or simply fun stuff! We are often brainstorming on initiative ideas and how to approach possible initiatives.

Aside from virtual interactions, we hope to establish a strong presence at the Joint Statistical Meetings (JSM) moving forward. We would like to (co)sponsor various sessions, as well as organize safe spaces for large and small group discussion, with and without facilitators.

We are also often in touch with our wonderful ASA Liaisons, Mark Otto, Regina Nuzzo, and Donna LaLonde.

What are some recent or upcoming committee events you would like to highlight?

Sarah: We are working hard to plan some exciting events and offerings for the ASA community. We're currently focusing on re-energizing this committee and announcing to the ASA community who we are and what we're about. We've collaborated with the JEDI and SPAIG committees to submit a proposal to co-sponsor a JSM 2022 session around the topic of inclusive collaborations. We are planning social media posts, podcasts, webinars, and a microsite that will host a variety of resources on statistics, disabilities, and accessibility for everyone to access.

Anja: We have been actively brainstorming since re-energizing the committee. We are a diverse group, so every member brings something different to the table in terms of initiatives.

An upcoming committee initiative that will be easily accessible to all members of the statistical practice will be our microsite. There, we plan to post content under various formats so microsite visitors of all backgrounds can consume the content.

For example, we have been talking about establishing a long-term virtual event initiative, which would offer a mix of longer webinar sessions (hosted live and recorded, subsequently posted to the microsite), as well as specialized capsules on a specific topic related to statistics, disability, and connected themes. The formats of the webinar sessions would also vary from didactic and topic-focused to a discussion panel or guided discussion by a facilitator.

By offering easily accessible content through the microsite, we hope to reach a wide audience and spread the word about the importance of accessibility and inclusivity in our community. We also hope to interact with the ASA and statistical community at large to better understand unfulfilled needs and how the CSD can contribute to bridging gaps.

Committee Members

Anja Zgodic, Committee Chair

Mark Otto

Matthew Brault

Erin Elizabeth Chapman

Charles Coleman

Peter Flom

Pamela McGovern

Ryan Machtmes

Sarah Zlatkovic

Maria Matilde Kam

Donna LaLonde (staff liaison)

What are some upcoming events you are most excited about?

Anja: Along with the SPAIG and JEDI committees, we look forward to hosting a session called "Making Data and Statistics Impactful Through Inclusivity, Powerful Collaboration, and Effective Innovation" at JSM 2022.

Our speaker will present on the topic of universal design (UD), which the University of Washington defines as "the process of creating products that are accessible to people with a wide range of abilities, disabilities, and other characteristics." The word "products" in this definition can refer to processes or workflows such as ones we would encounter in professional situations or academic settings, for example.

While objects can be designed according to UD, so can courses, conferences, content, even small things like fonts. Our speaker will cover the use of UD in academic departments to foster inclusivity and accessibility, which can lead to powerful collaborations resulting in innovation.

Any additional information or news you'd like to share about the committee with ASA members?

Sarah: If you are interested in getting involved with the Committee on Statistics and Disability, we'd love to welcome you as a Friend of the Committee! Friends can connect with the committee through Slack and attend meetings to hear what we're working on next. We welcome all interested to join the conversation.

Anja: We also hope to participate in joint initiatives with other ASA committees, interest groups, and ASA members. We invite any group interested in collaborating with us to get in touch. We also invite any member of the statistical community to join in our initiatives. ■

STAT*tr@k*

Advice on Applying for an Internship

Bill Pikounis and Elizabeth Mannshardt share their ideas and advice about the value of internships.

Should a person who satisfies most, but not all, of the listed qualifications apply for an internship?

Mannshardt: Absolutely. The list of qualifications listed on any job announcement is generally a 'wish list' for the employer. While there may be a couple of must haves, hiring managers are not necessarily looking for someone with every single qualification and, at the résumé review stage, are rather broadly looking for possible good matches. Additionally, there may be multiple opportunities available. Candidates with a subset of specific skills and qualifications are great matches for different options within the same hiring action. Multiple hires from one job announcement are relatively common, particularly for internships. It is also possible for résumés to be passed from one department to another. A hiring manager may recognize that what might not be a good fit for them may be a great fit for their colleague's team.

Pikounis: With proper expectations, yes, if the interest is strong. The one qualification I believe is negotiable with a prospective employer is the label of the major you are pursuing. Statistical internships no longer require only statistics majors. Any quantitative major should be considered, for example data science, informatics, etc.

What advice would you offer to address the missing qualifications?

Mannshardt: Focus on existing strengths and highlight relevant experience in your application materials. Be sure to include applicable job announcement wording in your résumé and cover letter. Missing qualifications can be addressed at the interview stage if the hiring manager inquires. You can point out similar skills (e.g., no direct Python programming experience but fluent in R/Julia over multiple types of projects) and discuss options for training and any necessary courses. Employers are looking for enthusiastic people who are willing to learn and tackle new projects.

Pikounis: Incorporate a description of your statistics-related skills and accomplishments in your CV / résumé. If interviewed for a position, be prepared to describe these.

Internship listings often list desired skills. What should an applicant do to communicate they have those desired skills?

Mannshardt: Highlighting relevant projects can be a way to showcase your experience and desired skills in a particular area. This can also serve to show collaborative work across scientific disciplines. List a couple of key tasks and/or milestones to which you heavily contributed or on which you regularly worked; describe any deliverable (article, report, presentation to management, etc.) Your résumé and cover letter should be updated for and tailored to each application. If you have worked on three main projects, they may all be listed on your résumé or CV, but consider emphasizing the most relevant project in your cover letter—possibly including a few technical details, your individual contributions, and what your learned or what experience you gained. Include applicable wording from the announcement in your own materials.

Pikounis: Concise, concrete, specific examples that demonstrate skills and accomplished results of value will strengthen the CV/résumé to be noticed favorably.

Is an online presence required? Should candidates have a digital portfolio, be on LinkedIn, etc.?

Mannshardt: It can, of course, be helpful to have an online presence but, in my experience, an extensive online presence is not required, particularly at the internship stage. In my experience hiring interns, initial screenings are done via the résumé and cover letter. There is often not time to do a detailed walk-through of each intern applicant's full set of online materials. However, skimming for relevant work may happen for top applicants within the résumé pool. This may be different across other industries and organizations. Other hiring managers may look extensively at applicants' digital portfolio and/or LinkedIn.

Pikounis: Not required but is nice to have, as long as it may add important information not present in the CV / résumé application.

In addition to the internship listings on STAT*tr@k*, are there other sources someone should check?

Mannshardt: There are many sources for job postings. Regularly checking ASA listservs and section and chapter websites should yield multiple options within specific topics and locations. For federal government employment, there are also multiple options. All federal jobs are posted on usajobs.gov, and there are additional sites specific to internships and new graduates, including Open Opportunities (*https://usajobs.github.io/openopps-help)*—government-wide program offering professional development opportunities to current federal employees and internships to students—and the Federal Pathways (*https://bit.ly/3kJPsWb*) program for federal internships. Federal job announcements are often posted for short periods of time—often five business days. I highly recommend those interested in federal employment set up a profile on USAJOBS. You can customize job search options that will notify you of postings that match the parameters you have set for each search—key words, job series (statistician, data scientist, IT, etc.), and location and/or agency/department.

Pikounis: Yes! Choose a company in an industry you are interested in and search their careers / job posting online site directly or use a search engine such as Google with company and industry terms and a phrase such as "summer internship." Along with advice to prepare your CV / résumé, check with your university resources for all potential channels.

"Speaking from personal experience, I learned just as much in my own internships as I did in any one specific course. Internships can also provide a chance to explore a particular field. Internships can help you determine what may (or may not!) be of interest to you—and one may even point you to a career path or opportunity you might not have otherwise considered."

– Elizabeth Mannshardt

Most internship applications require the submission of a CV. What advice do you have for creating and organizing a CV?

Mannshardt: How you organize your CV may be somewhat dependent on what types of positions you are applying for. For example, my academic CV lists my research projects and areas of interest before my job experience, whereas my nonacademic résumé puts my work experience and job titles at the top. Possible sections: degree (earned and upcoming); relevant work experience and past positions; key projects, including tasks/deliverables; relevant courses and specialized skills such as programming languages; volunteer work and professional service (e.g., ASA); awards; and publications.

What questions should a candidate ask to make sure the internship is a good fit?

Mannshardt: Who will you be directly working with? Will you have a mentor and guidance on day-to-day tasks, or will you be working largely independently? Ask about the team. Are you one of many, or will you be working on a solo project? Ask about what is of interest to you. Do you want to ML program all day? Do you want to work on one intense project for the entirety of your internship or are you interested in multiple projects? Are you looking for project management experience? Inquire as to the top priority on their side for the outcome of the internship. Is there a project that needs to be completed by a specific deadline? Are they doing general long-term research? What are they looking to get from their interns during the duration of the internship?

Pikounis: What will constitute a successful internship, from your perspective? What particularly stands out in your organization's culture? What projects do you foresee me working on? ■

Bill Pikounis (*www.linkedin.com/in/billpikounis*) is head of manufacturing and applied statistics at Janssen R&D, the pharmaceutical division of Johnson & Johnson. He oversees worldwide statistical services and products to serve areas of manufacturing of biotech (biologic, large), molecule, vaccine, and small molecule (pill, tablet) products. Pikounis has served as president of the Princeton-Trenton Chapter and Philadelphia Chapter and is an ASA Fellow.

Elizabeth Mannshardt (*www.linkedin.com/in/elizabeth-mannshardt-hawk-she-her-hers-6a1403169*) is the associate director of the US Environmental Protection Agency's Information Access and Analytic Services Division and an adjunct associate professor in the department of statistics at North Carolina State University. Mannshardt earned her PhD from The University of North Carolina at Chapel Hill. She is vice chair of the ASA's Committee on Career Development and also serves on the organizing committee for the North Carolina Chapter of Women in Machine Learning and Data Science.

2022 INTERNSHIPS

The following companies are looking for interns in 2022. If you are interested in bettering your programming techniques, data analysis skills, and software skills, apply for one of these opportunities.

MORE ONLINE
Find full descriptions for these internships on STAT*tr@k* at *https://stattrak.amstat.org/2021/12/01/2022-internships.*

If your organization would like to include an internship opportunity on our website, complete the form at *http://bit.ly/2CBAzyH*. Interested students will send a letter of inquiry and résumé directly to the contact and location you list.

AbbVie, Inc.

North Chicago, IL

Positions: Multiple

Type of Student: PhD candidates in statistics or biostatistics

Deadline: January 15, 2022

Requirements

- Be enrolled in a graduate-level curriculum leading to a PhD in statistics or biostatistics
- Have completed at least two full years of graduate study prior to the start of the internship
- Have an accumulated GPA of ≥ 3.0
- Be in good academic standing within your graduate program and overall at your university
- Have a track record of accomplishment
- Be authorized to work in the US
- Continue to be enrolled in graduate school for at least one semester following the internship

Apply: All candidates should apply on the AbbVie career website at *https://bit.ly/3kpflKA*.

Amgen

Thousand Oaks, CA

Positions: 1

Type of Student: Graduate, PhD

Deadline: January 31, 2022

Requirements

- Must be 18 years or older
- Graduated with a bachelor's degree from an accredited college or university with a 3.0 minimum GPA or equivalent
- Currently enrolled in a master's program from an accredited college or university and completion of the first year of master's program before internship starts
- Enrolled in an accredited college or university following the potential internship or co-op assignment
- Must not be employed at the time the internship starts

For information, visit *careers.amgen.com* and search for Keyword 129573.

Contact: Halle Dawson, *hdawson@amgen.com*

Apply: *http://adtrk.tw/tp/rj6_rqhp.y_K*

Astellas Pharma, Inc.

Northbrook, IL (remote possible)

Positions: 2+

Type of Student: PhD candidate in statistics or related discipline

Deadline: January 28, 2022

To be considered, applicants must have completed at least two years of graduate-level course work and be working on a dissertation toward a PhD in statistics or biostatistics. The applicant must be legally authorized to work in the United States. In addition, applicants should have a good working knowledge of R, S-Plus, or SAS and good communication skills.

Apply: Send CV, personal statement of interest, and a letter of recommendation to *Biostat.Intern@Astellas.com*.

AstraZeneca

Gaithersburg, MD; Durham, NC; Waltham, MA; Remotely

Positions: Multiple

Type of Student: PhD (preferred) or MS candidates in statistics or biostatistics

Deadline: February 28, 2022

AstraZeneca will have multiple full-time summer biostatistics internship openings for approximately 10–12 weeks. At the end of the internship, you will be expected to present your work within AZ Biometrics. Candidates must have a good knowledge of R and/or SAS.

Contact: Valerie Volpe, *valerie.volpe@astrazeneca.com*

Apply: *https://careers.astrazeneca.com/student-worker-and-internships-us* (select "apply now")

Auburn University

Auburn, AL

Positions: 4

Type of Student: Graduate

Deadline: January 10, 2022

This is an interdisciplinary program in which candidates accepted into Auburn's graduate program from an array of schools and departments, including math and statistics, can apply. Funded trainees receive a $34,000 stipend per year for up to two years, qualify for the Auburn graduate out-of-state tuition and instate tuition waivers, and have a set amount of semester student fees paid.

Candidates need to be accepted to the Auburn University graduate program before participating in the NRT program.

Contact: Kathryn Brown, *ksn0006@auburn.edu*

Apply: *https://bit.ly/2ZmtVei*

Biogen

Cambridge, MA

Positions: 7

Type of Student: PhD

Deadline: February 28, 2022

As an intern at Biogen, you can expect to be placed on a real project, under the guidance of experienced professionals and subject matter experts who are invested in your career and academic growth. We also ensure you have plenty of opportunities to build your network, learn about our organization through weekly lunch and learns led by leaders from across the company, and join us for several fun events.

Apply: *https://smrtr.io/6WGpW*

Bristol Myers Squibb

Lawrence Township, NJ

Positions: Multiple

Type of Student: Master's or doctoral students in statistics, biostatistics, or data sciences

Deadline: February 15, 2022

At Bristol Myers Squibb, we are inspired by a single vision: transforming patients' lives through science. In oncology, hematology, immunology, cardiovascular disease, and fibrosis—one of the most diverse and promising pipelines in the industry—each of our passionate colleagues contributes to innovations that drive meaningful change. We bring a human touch to every treatment we pioneer.

Contact: Ming Zhou, *ming.zhou@bms.com*

Apply: Data Science - *https://bit.ly/3EhUcJF*; Biostatistics - *https://bit.ly/3mg3BLF*

Daiichi Sankyo, Inc.

Basking Ridge, NJ

Positions: 3

Type of Student: PhD candidates in biostatistics or statistics

Deadline: February 15, 2022

Work closely with a senior-level biostatistician on statistical methodology and/or application topics related to the design and analysis of clinical trials in oncology, including Bayesian decision framework, safety-signal detection, literature review, and conduct simulation. Full-time internships in the summer will last 10–12 weeks. Starting and ending dates are flexible.

Contact: Minggao Shi, *mshi@dsi.com*

The D. E. Shaw Group

New York, NY

Positions: 4

Type of Student: Undergraduate, PhD

Deadline: March 1, 2022

We are the D. E. Shaw group, a global investment and technology development firm. *www.deshaw.com/careers/internship*

Eli Lilly and Company

Indianapolis, IN

Positions: Multiple

Type of Student: PhD

Deadline: January 15, 2022

Requirement: Must hold a PhD in statistics or biostatistics or have completed third year of PhD program by May 2022 and passed qualification exam

Therapeutic Area Knowledge
Understand disease states, competitive landscapes, and the regulatory environment to enhance the level of customer focus and collaboration and be seen as a strong scientific contributor

Regulatory Compliance
Perform work in full compliance with assigned curriculum(s) and be responsible for following applicable corporate, medical, local, and departmental policies, procedures, processes, and training

Statistical Leadership and Teamwork

- Introduce and apply innovative methodology and tools to solve critical problems
- Merge scientific thinking and business knowledge to identify issues, evaluate options, and implement solutions
- Lead projects independently and work effectively across functions
- Apply technical expertise to influence business decisions

Contact: Rachael Chickering, *Chickering_Rachael@Lilly.com*

Apply: *https://bit.ly/3jEWhaS*

Eli Lilly and Company

Indianapolis, IN

Positions: Multiple

Type of Student: Master's

Deadline: January 15, 2022

Requirement: Must hold master's in statistics or biostatistics or have completed third year of PhD program by August 2023 and passed qualification exam

Therapeutic Area Knowledge
Understand disease states, competitive landscapes, and regulatory environment to enhance customer focus and collaboration and be seen as a strong scientific contributor

MORE ONLINE
Find full descriptions for these internships on STAT*tr@k* at *https://stattrak.amstat.org/2021/12/01/2022-internships.*

Regulatory Compliance
Perform work in full compliance with assigned curriculum(s) and be responsible for following applicable corporate, medical, local, and departmental policies, procedures, processes, and training

Statistical Leadership and Teamwork

- Introduce and apply innovative methodology and tools to solve critical problems
- Merge scientific thinking and business knowledge to identify issues, evaluate options, and implement solutions
- Lead projects independently and work effectively across functions
- Apply technical expertise to influence business decisions.

Contact: Rachael Chickering, *Chickering_rachael@lilly.com*

Apply: *https://bit.ly/3mdImdA*

Genentech, Inc.

South San Francisco, CA

Positions: Multiple

Type of Student: Master's, PhD

Deadline: January 14, 2022

The Data and Statistical Sciences summer interns will work for 10–12 weeks on theoretical or applied problems with direct relevance to ongoing clinical or nonclinical drug development in diverse therapeutic areas such as oncology, immunology, infectious disease, ophthalmology, and neuroscience. The specific topics may cover research problems from translation research to late-stage clinical trials, new methodologies in machine learning to enable scientific reverse translation using clinical trial and real-world data, and innovative solutions for automating data handling and analytics processes.

Contact: Cristian Oliva-Aviles, *olivaavc@gene.com*

Apply: *go.gene.com/dss-summerintern*

GlaxoSmithKline

Upper Providence, PA

Positions: 2

Type of Student: PhD

Deadline: January 28, 2022

The selected candidates will have the opportunity with other statisticians and GSK scientists to address drug discovery questions using statistical methods such as experimental design, mixed models, repeated measures, analysis of high-dimensional data, machine learning algorithms, and Bayesian methods. The internships may be onsite or remote.

Candidates must be currently enrolled students seeking their PhD in statistics, biostatistics, or a closely related field and must have completed two years of graduate study (i.e., coursework equivalent to an MA/MS in statistics).

Contact: *andrew.x.gehman@gsk.com*

Apply:
www.gskbiostatisticscareers.com

Google

United States and Canada

Positions: n/a

Type of Student: PhD

Deadline: January 28, 2022

Work with large complex data sets, solve difficult nonroutine analysis problems, and apply advanced analytical methods as needed

Research and develop analysis, forecasting, and optimization methods to improve the quality of Google's user-facing products such as ads quality, search quality, end-user behavioral modeling, and live experiment

Requirements

- Pursuing a PhD in statistics, biostatistics, operations research, physics, economics, applied mathematics, or similar quantitative discipline
- Experience with statistical software (R, Python, S-Plus, SAS, or similar) and database languages (e.g., SQL)
- Experience with data analysis

Apply: *https://bit.ly/2XKXzt2*

Institute for Defense Analyses (IDA)

Alexandria, VA; Washington, DC

Positions: 15–20

Type of Student: Undergraduate (rising seniors) and graduate students

Deadline: January 18, 2022

All current and newly hired IDA employees must be fully vaccinated against COVID-19 unless they have been granted a medical or religious exemption.

The IDA Summer Associate Program provides students a unique opportunity to use their quantitative skills to work on national security issues.

Each year, IDA hires paid student associates at its Virginia and Washington, DC, facilities to work for a minimum of 10 weeks in the summer. Attendance at the eight-week core period between June 6, 2022, and July 29, 2022, is required. An additional two weeks either before or after the core period is required for a complete 10-week internship.

Qualifications
We seek qualified students with strong academic backgrounds.

Our preference for the program is graduate students, but we will consider undergraduates who are rising seniors.

SAC seeks a variety of disciplines, including computer science, information technology, mathematics, chemistry, physics, operations research, aero/astronautical engineering, mechanical or electrical engineering, materials science, statistics, economics, social and

behavioral sciences, human resources, security, international relations, policy studies, and finance. We are also recruiting students focusing on library and information science, English, history, and related areas.

STPI seeks a variety of disciplines, including behavioral sciences, biological sciences, computer science, economics, engineering, environmental sciences, learning sciences, mathematics, physical sciences, social sciences, and statistics.

Applicants are required to submit a résumé and transcripts.

Applicants who are selected will be subject to a security investigation.

Criminal background check is required.

Apply: *https://bit.ly/3vLpfuo*

Johnson & Johnson

Spring House, PA; Titusville, NJ; Raritan, NJ; La Jolla, CA

Positions: 10

Type of Student: PhD candidate

Deadline: February 15, 2022

In this role, students will be involved in activities such as supporting research and development in areas that range from drug discovery through phase 2 and 3 clinical studies and, concurrently, pharmaceutical manufacturing, including data preparation, graphical exploration of data, statistical model building and analyses, report writing, and writing of computer programs or applications associated with the statistical analyses.

Qualifications and Notes

- Candidates must be enrolled in an accredited college (not necessarily taking classes) and pursuing a PhD in biostatistics, statistics, data science, or a related discipline.
- Students must be available for 10–12 weeks from May to August and have the ability to work full-time during that time.
- A minimum 3.0 GPA is preferred.
- Remote work options will be available. If J&J onsite internships are permitted by public health governance in summer 2022, students must be able to provide their own transportation to/from their work location.

In December, the JNJ careers site will have a general posting for internships as described above and will include the ability to apply. The search terms "intern statistics" or "SDS" (without the quote marks) should provide a listing that includes the summer intern posting to review and apply.

The Lubrizol Corporation

Wickliffe, OH

Positions: 4

Type of Student: Undergraduate, graduate, PhD

Deadline: February 11, 2022

The statistics and data analytics team is charged with creating systems that enable highly effective product development via virtual experimentation, optimization, and knowledge discovery. In addition, the team provides data science consulting services to the Lubrizol technical community. The project work depends on the skills/interests of the intern and the current needs of the department.

Contact: Jessica Schmiesing, *Jessica.Schmiesing@Lubrizol.com*

Apply: *lubrizol.com/careers*

Novartis

Remotely; East Hanover, NJ; Cambridge, MA; Fort Worth, TX; Princeton, NJ; San Diego, CA

Positions: Multiple

Type of Student: Graduate, PhD

Deadline: January 21, 2022

Candidates must be enrolled in a graduate-level program working toward a PhD in biostatistics, statistics, pharmacometrics, computer science, engineering, or a related discipline and have completed at least 1.5 years of course work. Competitive candidates must have excellent oral and written communication skills and strong problem-solving skills. Working knowledge of R or SAS—as well as a strong background in NONMEM, Python, and/or other software/languages—is preferred.

Apply: Send your CV to *https://forms.office.com/r/Zv32QtBLAb.*

Netflix

Los Gatos, CA

Positions: 4

Type of Student: PhD

Deadline: January 17, 2022

Requirements

- Pursuing a PhD, entering third year or later in fall 2022, at an accredited university and authorized to work as a summer intern in the USA
- Engaged in relevant research into online experimentation, sequential experimentation, adaptive experimentation, causal inference, or related area
- Experience programming in R and/or Python
- Curious, self-motivated, and excited about solving open-ended challenges at Netflix
- Great communication skills, both oral and written

Internships are paid and are a minimum of 12 weeks, with a choice of a few fixed start dates in May or June to accommodate varying school calendars. Conditions permitting, our summer internships in 2022 will be located—depending on the team—either remotely, in our Los Gatos office, or in our Los Angeles office.

Apply: *https://jobs.netflix.com/jobs/122370775*

MORE ONLINE
Find full descriptions for these internships on STAT*tr@k* at *https://stattrak.amstat.org/2021/12/01/2022-internships.*

Office of Biostatistics, Center for Drug Evaluation and Research, US Food and Drug Administration

Silver Spring, MD

Positions: Multiple

Type of Student: Graduate students with strong background coursework in biostatistics or statistics; completion of doctoral prequalifying exams preferred

Deadline: March 1, 2022

We anticipate having multiple internship positions available for advanced PhD graduate students in statistics or biostatistics from May 16, through September 2, 2022, to engage in research projects on topics relevant to the Office of Biostatistics (OB) scientific needs.

Interns are expected to participate 40 hours per week at our headquarters in Silver Spring, Maryland, with their mentor and cohort.

Participants may be able to conduct offsite training on a temporary basis while the mentor and FDA staff work offsite. Interns will prepare a written report and a 30-minute presentation.

Requirements

- Excellent oral and written communication skills
- Interpersonal and teamwork skills
- Strong problem-solving skills
- Strong computational skills
- Self-management skills
- Solid statistical background
- Proficiency with MS Office
- Experience with SAS and/or R
- Demonstrated creativity

Apply: Send CV and cover letter to *CDER-OTS-OB-Recruitment@fda.hhs.gov* with APPLICATION ORISE 2022 in the subject line.

Questions: Visit *https://bit.ly/306CYjM* or email *CDER-OTS-OB-Recruitment@fda.hhs.gov*; use QUESTION ORISE 2022 as the subject.

MORE ONLINE
Find full descriptions for these internships on STAT*tr@k* at *https://stattrak.amstat.org/2021/12/01/2022-internships.*

Pfizer Inc.

La Jolla, CA; Boulder, CO; Groton, CT; Collegeville, PA; Cambridge, MA; New York, NY; Andover, MA; Pearl River, NY; Remotely

Positions: Multiple

Type of Student: Graduate students in statistics, biostatistics, or related fields

Deadline: January 31, 2022

The internship will consist of up to 480 hours of work at one of the Pfizer sites, starting in early April and ending in December. The intern's project will be biomedically oriented, with one-on-one supervision by one or multiple senior staff statisticians. The work will be a hands-on learning focusing on current project needs and will likely involve use of SAS, and R. As part of the program, the intern will prepare a written report and brief presentation summarizing the work.

Applicants must have completed at least one year of graduate study by May 2022 and have work authorization in the US.

Apply: Submit your application via Handshake (Job ID: # 5512484 - 2022 Pfizer Statistics Internships). Visit *https://app.joinhandshake.com/jobs/5512484.* If you do not have a Handshake account, you will be prompted to create one.

QSURE at Memorial Sloan Kettering Cancer Center

New York, NY

Positions: 10

Type of Student: Undergraduate

Deadline: January 12, 2022

The QSURE internship program is designed for motivated undergraduate students with a passion for quantitative science. Each student will be paired with an expert mentor to work on an individual, data-driven research project that aims to answer pressing questions about cancer. The program will provide hands-on skills in biostatistics, epidemiology, health outcomes, and/or computational oncology.

QSURE interns will receive training in scientific and career development skills, including the following:

- Hands-on research experience carried out under the guidance of a faculty mentor
- Quantitative skills and familiarity with methods used in cancer and population health research
- Exposure to a wide application of data through internal and external speakers
- Education in the responsible conduct of research
- Training in oral and written scientific presentation
- Resources for career development
- Understanding of options for graduate study and careers
- A supportive cohort of like-minded peers

A modest stipend will be provided.

Applicants must be authorized to work in the US. Students are responsible for their own housing, but QSURE will provide a nominal reimbursement for housing expenses.

Contact: *bstQSURE@mskcc.org*

Apply: *mskcc.org/qsure*

Sanofi US, Inc.

Bridgewater, NJ; Boston, MA

Positions: Multiple

Type of Student: PhD candidate in statistics or biostatistics

Deadline: March 1, 2022

Successful candidates will work on the design and analysis of early- and late-phase clinical trials and statistical methodology research under the supervision of senior-level statisticians.

Candidates must have completed at least two years of graduate coursework and be working on a dissertation toward a PhD in statistics or biostatistics. Requirements include

effective oral and written communication skills and knowledge of SAS and/or R. Python is a plus.

Contact: Xiaodong Luo, *xiaodong.luo@sanofi.com*

Apply: Email CV and (un)official graduate transcript to *xiaodong.luo@sanofi.com.*

Science Undergraduate Laboratory Internship (SULI) Program Community College Internship (CCI) Program at Pacific Northwest National Laboratory

Remotely

Positions: 50+

Type of Student: Undergraduate, Community College

Deadline: January 12, 2022

The Science Undergraduate Laboratory Internships (SULI) program encourages undergraduate students and recent graduates to pursue STEM careers by providing research experiences at the Department of Energy (DOE) laboratories. Selected students participate as interns at one of 17 DOE facilities. They perform research under the guidance of laboratory staff scientists or engineers on projects supporting the DOE mission.

The Community College Internships (CCI) program seeks to encourage community college students to enter technical careers relevant to the DOE mission by providing technical training experiences at the DOE laboratories. Selected students participate as interns appointed at one of 16 participating laboratories. They work on instrumentation projects or major research facilities supporting DOE's mission under the guidance of staff, scientists, or engineers.

Contacts: Nicole Castilleja Bentley, *nicole.bentley@pnnl.gov*; Nancy Roe, *nancy.roe@pnnl.gov*

Thomas Jefferson University, Division of Biostatistics

Philadelphia, PA

Positions: 3

Type of Student: Undergraduate (junior or senior preferred); graduate (MS or PhD)

Deadline: February 15, 2022

Jefferson's Division of Biostatistics will sponsor up to three students as interns, who will do the following:

- Research statistical topics relevant to biomedical research
- Apply statistical thinking to biomedical research problems
- Analyze real-world biomedical data and interpret the results
- Develop statistical programming skills in SAS, R, and other languages
- Practice communication of statistical methods and results through written reports and oral presentations
- Receive guidance regarding their future studies and career trajectory

The internship will run from June to August. Depending on the COVID-19 situation, it will take place either in person or remotely. Interns will be paid a stipend.

Applicants may be current students or recent graduates and have an interest in pursuing a career in (bio) statistics, data science, or other similar quantitative field.

For undergraduate students (or those with an undergraduate degree), any degree program is acceptable, but applicants should have strong mathematical and analytical skills and at least one course in computer/statistical programming.

For graduate students (or those holding a graduate degree), their degree should be in biostatistics, statistics, computer science, or a similar quantitative field.

Applications will be reviewed and requests for interviews will be sent on a rolling basis.

Contact: Constantine Daskalakis, *constantine.daskalakis@jefferson.edu*

Apply: *https://jeffline.jefferson.edu/education/programs/biostatistics_si*

Travelers

Hartford, CT; St. Paul, MN

Positions: 35

Type of Student: Graduate (MS and PhD)

Deadline: February 15, 2022

Join us at Travelers as a Data Science Leadership Development Program (DSLDP) intern! As a DSLDP intern, you will be exposed to a challenging professional work experience in data science with planned activities giving you insight into the community at Travelers and the many ways we solve business problems using cutting-edge statistical and analytical techniques. You will assist in design and develop and program methods, processes, and systems to consolidate and analyze unstructured and structured diverse data sources to generate actionable insights and solutions for client services and product enhancement.

With many company-wide events such as intern symposium and community outreach events, you will have the opportunity to interact and network with many levels of leadership and management. You will be expected to work with fellow interns to develop and present solutions to relevant business and data science topics.

The internship offers a training curriculum focused on data science and business acumen. Other program components include formal mentoring, networking, and career guidance.

Contact: Alicia Hitchcock, *AHITCHCO@travelers.com*

Apply: *https://travl.rs/3C5S7iq* ■

MORE ONLINE
Find full descriptions for these internships on STAT*tr@k* at *https://stattrak.amstat.org/2021/12/01/2022-internships.*

JSM 2022 AUGUST 6–11 WASHINGTON, DC

Call for Contributed Abstracts, Chairs

Ming-Hui Chen, JSM 2022 Program Chair

MORE ONLINE
Submit your abstract today! *https://ww2.amstat.org/meetings/jsm/2022/beontheprogram.cfm*

JSM 2022 will be held in Washington, DC, August 6–11. This is going to be the largest in-person statistical conference held in North America since the COVID-19 pandemic started. The theme for JSM 2022 is "Statistics: A Foundation for Innovation," which emphasizes the essential role of statistics in the era of big data and data science.

The JSM 2022 Program Committee has put together 164 invited sessions, including both paper and panel sessions. The topics of panel sessions cover areas such as teaching innovations, career development, real-world evidence in drug development, promoting diversity, the role of statistics in science and policymaking, and effective reporting of statistical results. Like previous JSMs, the topics of paper sessions are interesting and diverse, including modern statistical learning methods, precision medicine, microbiome research, causal inference, innovative clinical trial design, and advanced statistical methods.

As always, many strong invited and topic-contributed session proposals were not selected due to the extremely competitive selection process. There are still ways to get involved in the 2022 JSM program, however.

Speed Sessions

The speed session is becoming increasingly popular at JSM. It allows for an electronic poster (e-poster) presentation, which enables video and other special effects. A speed session will consist of 20 oral presentations of approximately four minutes, with a five-minute break after the first set of 10 talks. These short oral presentations will be followed by an e-poster session later. Each speed poster session will last 45-minutes.

Only for the speed sessions will the regular 110-minute contributed poster sessions be divided into two sessions. There will be 45-minutes for a first group of 20 presenters, a 20-minute transition period, and then 45-minutes for a second group of 20 presenters. The program committee tries to cluster speed session posters by topic to attract a large and focused audience.

The following incentives will be offered to presenters who participate in speed sessions:

- Electronic poster boards, so there will be no additional costs or hassle associated with printing or transporting a poster
- Ability to present orally and through an electronic poster

Here are some tips, based on experience with previous speed sessions:

- The oral component should lure people. Don't try to be too detailed, but rather give the big picture view. A little humor helps.
- E-posters can include software demonstrations, analysis animations, videos, and interactive statistical graphics or dashboards. Take advantage of the versatility of the medium. Don't think in terms of a static poster. Be modern and daring.

When you submit your contributed abstract, simply select "Speed" as the sub-type.

Poster Sessions

JSM 2022 is expected to take place in person. Poster sessions permit face-to-face extended discussions with individuals or small groups interested in your topic. Particular advantages are direct feedback and the ability to display extensive graphical or tabular materials, possibly in addition to a handout.

Contributed Sessions

Nearly half of JSM sessions are contributed sessions. Contributed paper sessions consist of seven papers with 15 minutes of presentation time for each, including the introduction of the speaker and questions. Contributed abstracts submission closes February 1, 2022, and a decision about acceptance will be made on March 31, 2022.

Session Chairs

Each JSM session requires a chair. The responsibilities include contacting speakers with session information before JSM and introducing speakers and managing presentation time during the session. Chairing a session is a great way for researchers who are new to the profession to build a professional network and get involved with JSM (and one can mention this service on one's CV). Simply volunteer to the program committee members of your section or society. Contact JSM 2022 Program Chair Ming-Hui Chen at *ming-hui.chen@uconn.edu* with any questions. ■

Goals Met, Celebrated at Women in Statistics and Data Science Conference

Donna LaLonde, ASA Director of Strategic Initiatives and Outreach

Once again, those attending the 2021 Women in Statistics and Data Science (WSDS) Conference realized their goals of celebrating successes, growing influence, building community, and sharing knowledge.

The Caucus for Women in Statistics (CWS) and Association for Women in Mathematics (AWM) were founded in 1971, so WSDS 2021 was an opportunity to join their ongoing 50th anniversary activities. Former CWS presidents—**Nancy Flournoy, Nancy Gordon, Stephanie Shipp,** and **Jiayang Sun**—joined current president **Tomi Mori** and executive director **Jessica Kohlschmidt** for a plenary panel to reflect on the 50-year history of CWS and look to the future.

Mary Gray, who is the founder of AWM and a former president of CWS gave a plenary talk, "Leave No Data Unturned." Throughout her career, Gray has been an advocate for social justice. She told us, "If you want to help, there are always interesting places to do so. I found myself teaching regression in Fiji, studying school finance in Myanmar, and returning a few times to Iraq."

These anniversaries were complemented by a plenary presentation celebrating a new initiative. **Talitha Washington,** who is the inaugural director of the Atlanta University Center Data Science Initiative and a professor of mathematics at Clark Atlanta University, presented a plenary session titled, "Strategic Leadership in the Pursuit of Data Science for Black Lives." We learned the new initiative is to provide data-driven solutions to current and emerging societal problems, especially as it pertains to the African American community. We also learned about her outlook on leadership.

Many concurrent sessions continued the WSDS tradition of recognizing both challenges and opportunities.

In the session It's Never Too Early to Develop Leadership Skills, the following presenters discussed first steps into leadership:

- **Alexandra L. Hanlon,** Center for Biostatistics and Health Data Science, iTHRIV Biostatistics, and Virginia Tech
- **Nichole E. Carlson,** Center for Innovative Design and Analysis, Colorado School of Public Health
- **Leslie McClure,** Drexel University
- **Dionne Price,** Division of Biometrics IV, Office of Translational Sciences, Center for Drug Evaluations and Research, Food and Drug Administration
- **Sally C. Morton,** Arizona State University

In the panel session Mindful Networking to Advance Diversity, Equity, and Inclusion, **Emma Grace Thomas** of RAND Corporation engaged colleagues **Emma Benn** of the Icahn School of Medicine at Mount Sinai, **Heather Mattie** of Harvard T.H. Chan School of Public Health, and **Natalie Cheung Rotelli** of Eli Lilly and Company in a conversation about how to build and leverage diverse professional networks.

In a session recognizing the Caucus for Women in Statistics inaugural Societal Impact Award, recipient **Melody S. Goodman** and her colleague **Rebecca Betensky** spoke about the Quantitative Public Health Data Literacy Training Program and the pipelines into the Quantitative Aging Research Summer Program.

These sessions are examples of the breadth of the WSDS 2021 conference program.

The 2021 conference was supposed to take place in Pittsburgh, Pennsylvania, but was changed to a virtual conference due to COVID-19. Although the format changed, the goals of celebrating successes, growing influence, building community, and sharing knowledge remained and were met. ■

Meet Us in St. Louis

Did you miss this year's meeting? Join us next year in St. Louis, Missouri, October 6–8. Submission will be open for concurrent, panel, and poster session abstracts from February 10 – April 15. For details and to sign up for the WSDS mailing list, visit *ww2.amstat.org/meetings/wsds/2022.*

MORE ONLINE
Visit the ASA's Flickr account to see screen shots of the online event: *https://bit.ly/3n3vyXJ.*

ASA Fellow **Joe Cappelleri,** head of outcome research statistics, will receive this year's International Society of Pharmacoeconomics and Outcomes Research (ISPOR) Avedis Donabedian Outcomes Research Lifetime Achievement Award for his outstanding contributions over the years to patient health outcomes.

Cappelleri is the first pharmaceutical industry statistician to receive this award. In a letter to the review committee, Brown University adjunct assistant professor Jessica Roydhouse, wrote, "Joe's impressive leadership in patient-reported outcomes has led to the development and psychometric validation of widely used tools in different disease areas. These tools have been included in clinical trials as part of regulatory evaluation and facilitated the inclusion of the patient experience in drug labels."

Cappelleri has published approximately 600 articles, many in leading medical and statistical journals, and has been cited more than 24,000 times; 64 of his publications have been cited at least 64 times and 210 of them have been cited at least 10 times.

The ISPOR Avedis Donabedian Outcomes Research Lifetime Achievement Award recognizes an individual's outstanding, life-long achievement in improving health outcomes. The award is international in scope and stature.

The award was established in honor of the late Avedis Donabedian, who has been called the "father of outcomes research." Donabedian was a renowned faculty member at the University of Michigan School of Public Health, multiple award-winner, author of eight books and more than 50 peer-reviewed articles, and presenter of countless lectures on outcomes research. He dedicated his life to improving the quality of health care and health care systems, directing such research toward health outcomes as the measure of quality. ■

Imbens

ASA Fellow **Guido W. Imbens,** Joshua D. Angrist, and David Card were awarded the 2021 Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2021. David Card earned half the winnings "for his empirical contributions to labour economics" and Angrist and Imbens split the other half "for their methodological contributions to the analysis of causal relationships." Imbens's seminal piece with Angrist was published in the *Journal of the American Statistical Association* in 1996.

Read more about this year's laureates on the Nobel Prize website, *https://bit.ly/3n5LicL*. For more information about the prize, visit *www.nobelprize.org*. ■

ASA Fellow **Cynthia Rudin** recently won the $1 million Squirrel AI Award for Artificial Intelligence for the Benefit of Humanity from the Association for the Advancement of Artificial Intelligence.

Rudin, former chair of the Statistical Learning and Data Science Section, is being cited for "pioneering scientific work in the area of interpretable and transparent AI systems in real-world deployments, the advocacy for these features in highly sensitive areas such as social justice and medical diagnosis, and serving as a role model for researchers and practitioners."

Currently a professor at Duke University, Rudin earned undergraduate degrees in mathematical physics and music theory from the University at Buffalo before completing her PhD in applied and computational mathematics at Princeton. She is also a three-time recipient of the INFORMS Innovative Applications in Analytics Award, which recognizes creative and unique applications of analytical techniques.

Read more about the prize and Rudin's work at *https://bit.ly/3n5L2dN*. ■

Sherri Rose, associate professor of health policy at the Stanford School of Medicine and co-director of the Health Policy Data Science Lab was awarded the 2021 Mortimer Spiegelman Award. The award recognizes a statistician under age 40 who has made the most significant contributions to public health statistics. Rose's research has been featured in *The New York Times, USA Today,* and *The Boston Globe.*

Rose earned her PhD in biostatistics from the University of California, Berkeley, and a BS in statistics from The George Washington University before completing an NSF Mathematical Sciences Postdoctoral Research Fellowship at The Johns Hopkins University. Prior to joining the faculty at Stanford University, Rose was on the faculty at Harvard Medical School in the department of health care policy. Read more about Rose and her work at *https://healthpolicy.fsi.stanford.edu/people/sherri-rose*. Learn about the Health Policy Data Science Lab at *http://healthpolicydatascience.org*. ■

Statistical Learning and Data Science Has Year of Successful Webinars

Jaime Lynn Speiser, Assistant Professor of Biostatistics and Data Science, Wake Forest School of Medicine

The Section on Statistical Learning and Data Science launched a monthly webinar series in 2021. The webinars, free for attendees, highlighted topics in data science across industry and academia. Each webinar included both introductory and advanced material.

Invited speakers were selected based on their expertise in mathematics, statistics, biostatistics, computer science, and industry. They represented a diverse pool of experts at differing levels in their careers, each with a unique perspective to share.

The webinars took place on Zoom toward the end of each month. Attendees registered for each webinar through a link shared on the section's discussion board. In total, the section sponsored 11 webinars in 2021. Four main themes emerged: career advice; reproducibility; interpretability; and methods in data science.

The first webinar theme was career advice for success in data science. **Helen Zhang** spoke about building data science teams based on her experience with the University of Arizona Transdisciplinary Research in the Principles of Data Science network. **Wayne Lee** shared his experiences across industry and gave advice about careers in data science in classical industries (e.g., agriculture and manufacturing) versus modern industries (e.g., technology). **Sarah Kalicin,** the section's president, gave a webinar about successful career relationships involving mentoring, coaching, and sponsorship. All the speakers gave advice about communicating within data science teams, seeking different types of careers in industry and academia, and leveraging relationships for career advancement.

Two webinars focused on reproducibility in data science. **Brian Lee Yung Rowe** provided a framework for reproducibility and automation in data science projects based on computational graphs. **Byron Jaeger** gave an overview of Git and GitHub that facilitate easy version control and collaboration. Given the current reproducibility crisis in our field, these webinars were timely and informative for improving reproducibility in data science projects across industry and academia.

The next theme emerging from the series was interpretable data science and artificial intelligence. **Beth Wolf** discussed variable importance measures for common machine learning models with examples in medicine. **Polo Chau** shared projects involving scalable, secure, and interpretable artificial intelligence with applications in image recognition and cyber security. These webinars highlighted the importance of interpretability in data science and understanding inputs and outputs of data science models.

The final theme for the series was advancing data science methods. **Naomi Brownstein** presented a testing framework for clusterability that allows for determining if unlabeled data has a cluster structure. **Andreas Ziegler** discussed calibration techniques for binary outcome models that can be used to adjust predictions and increase accuracy for validation with external data. **Jean Feng** gave an overview of deep learning and shared her work involving methods for variable selection with deep learning for small-scale data. **Nesime Tatbul** presented about challenges and opportunities for time series analysis with data science and introduced new methods for anomaly detection.

These four webinars feature novel methods being developed in data science. A take-away is that there are many opportunities for novel methods development in data science.

The section's first year of monthly webinars was a success. Webinar recordings are available at *https://community.amstat.org/slds/meetings/upcomingevents.*

If you have suggestions for webinar topics and /or speakers, email *sldswebinar@gmail.com.* ■

MORE ONLINE
Webinar recordings are available at *https://community.amstat.org/slds/meetings/upcomingevents.*

Survey Research Methods

New officers for 2022 are Brady West (chair-elect), Andreea Erciulescu (program chair), Qixuan Chen (secretary), and Kristen Olson (COS representative).

Outgoing officers include Morgan Earp (past chair), Brady West (program chair), Gina Walejko (secretary), and Julia Soulakova (COS representative).

MORE ONLINE
To view the complete section news, visit *https://magazine.amstat.org/blog/category/membernews/amstatsections.*

Nine SRMS members were named ASA Fellows this year. They are Julia Bienias, Stephanie Eckman, Elena Erosheva, Stanislav Kolenikov, Yan Li, Bo Lu, Pamela McGovern, Sowmya Rao, and Jasjeet Sakhon.

All SRMS members can access the 2011–2020 AAPOR webinars for free by emailing Brady West at *bwest@umich.edu.* ■

Biometrics

Esther Drill, principal biostatistician at Memorial Sloan Kettering Cancer Center, is the recipient of the 2020 ASA Biometrics Strategic Initiatives Grant. Her project, "Developing the Next Generation of Biostatisticians," in alignment with the "Bridge to Biostats" program, proposed a "Biostat days" to introduce underserved and underrepresented high-school students to biostatistics. Here, Drill provides an overview of the project and discusses the challenges and progress on the impact of biostatistics outreach amid the COVID-19 pandemic.

As chair of Memorial Sloan Kettering's Bridge to Biostats Committee, I proposed to develop a suite of interactive "statistical thinking" exercises for use as part of our Biostats Day awareness program, which introduces STEM-interested New York City high-school students from underrepresented minority (URM) groups to the biostatistics field. We present this hour-long program over Zoom to students from existing NYC-area organizations that serve our target audience. The proposed statistical thinking activities summarize and visualize data from students in real time, which allows for active participation in learning about introductory statistical concepts, including sampling, sample mean, and sources of error.

To date, we have completed the "Population and Sample Mean" activity, which compares the CDC's 45,000-person sample mean height with our student group's mean height, and "The F-Test" activity, in which all students have two minutes to count the number of Fs in the same paragraph and are presented with a histogram of the results demonstrating measurement error.

We have included these activities in Biostats Day presentations to 70 students from three NYC-based STEM exposure organizations: BEAM NYC (Bridge to Enter Advanced Mathematics), the Einstein Enrichment Program (a New York State–funded Science Technology Entry Program), and Memorial Sloan Kettering's Summer Exposure Program (SEP). While only 26 percent of students had heard of biostatistics before our presentation, 80 percent expressed interest in learning more about biostatistics at the end. The interactive activities were one of the most frequent answers to the question, "What is your favorite part of Biostats Day?"

Our third statistical thinking activity explores other aspects of error in the context of aging estimation; it is under active R Shiny development. We expect the "Guessing Ages" activity to be ready for use in our next Biostats Day presentation, as well as our BEAM Saturday enrichment class in Spring 2022 and our second cohort of SEP biostatistics summer students in 2022.

We are in the process of creating a GitHub repository for all our Biostats Day materials, including access to and instructions for the statistical activities, and hope it will be a valuable resource for groups interested in performing their own outreach."

View all the section's news online at *https://magazine.amstat.org/blog/category/membernews/amstatsections/biometrics.* ■

Quality and Productivity

Xinwei Deng, Section Chair-Elect

The Quality and Productivity (Q&P) Section executive committee met October 29 to discuss areas of focus for 2022. The committee discussed areas of strength, areas of weakness, and areas needing special attention.

Q&P is financially healthy, with savings growth from year to year. However, more effort is needed to keep documentation up to date (e.g., list of members, agenda), ensuring a smooth transition from current officers to elected officers. Additionally, significant improvement has been made over the past two years on updating the operation manuals and documentation on the roles of Q&P officers. However, a systematic approach can be taken with a more organized repository. The following are five focus areas for 2022:

1 – Refine the Q&P section charter and operating manual, making them concise and easy to access.

2 – Create a Q&P calendar (i.e., key date road map) to mark important action items. For better communication and transparency, investigate how to create the Q&P calendar to clarify Q&P officers' duties and remind them about planning and activities in advance.

3 – Create a cloud-based sharing folder (e.g., Google drive, GitHub) to synchronize documents and meeting notes.

4 – Improve the transition of officer positions, such as treasurer. With Q&PRC and FTC meetings under Q&P, there is an emerging need for creating an additional position (such as assistant treasurer) to make the transition smoother.

5 – Increase membership by advocating Q&P membership in related conference venues such as Q&PRC, FTC, and SRC. It is important to attract young professionals and statisticians to join Q&P. Conference venues such as Q&PRC, FTC, and SRC are good for recruiting new members. ■

Bay Area Chapter Hosts Triumphant Summer Project Program

Anwen Huang, Elaina Li, Avelyn Liang, Anna Khodakovskaia, Katherine Tsvirkunova, Vivian Wang, and Ron Yu

For the third year, the ASA's Bay Area Chapter (SFASA) hosted its K–12 student summer project program. The program brings together students to design projects under the mentorship of SFASA officers and volunteers.

This year's program began with a virtual kick-off meeting on June 4. Students were assigned mentors and brainstormed research topics together, providing feedback on everyone's ideas. Over the summer, the students worked on their projects and met with their mentors weekly to discuss obtaining the data and appropriate data analysis.

In contrast to previous years' survey-based projects, this year's students analyzed publicly available large data sets by using RStudio. Students learned to use RStudio from online tutorials and performing statistical analyses under their mentors' guides. On September 12, the students presented their projects virtually via Zoom. The presentations were open to all SFASA members; roughly 30 attended.

Elaina Li (8th grader) and **Avelyn Liang** (10th grader) presented their project on COVID-19 vaccination rates in California and their variation based on demographic characteristics and social-economic factors. Under the mentorship of **Ray Lin** (president of SFASA/Genentech) and **Priscilla Yen** (Amgen), they found statistically significant differences in vaccination rates across age groups but not across racial groups. Further, they discovered a heterogeneous relationship between vaccination rate and social factors such as median household income within both age and race groups.

Anna Khodakovskaia (8th grader) and **Katherine Tsvirkunova** (12th grader) analyzed beach clean-up data in the US under the mentorship of **Tao He** (past president of SFASA/San Francisco State University). They focused on the amount of plastic found on US beaches from states along the West Coast and East Coast and found the West Coast had a higher proportion of plastic in beach trash. They also determined insufficient evidence to assert that the larger the population, the greater trash collected per mile.

Anwen Huang (12th grader) explored the use of statistical models in breast cancer prediction under the guidance of **Ron Yu** (public relations director of SFASA/Gilead Sciences). Using RStudio, she compared the misclassification rate of a classification tree, logistic regression, and support vector machine models on the Wisconsin Breast Cancer Data Set, which recorded 10 variable values for malignant and benign breast tumor cells. The rates were calculated through five-fold cross-validation, and Huang determined a support vector machine with six or seven variables produced the highest classification accuracy of 95.1 percent.

Vivian Wang (10th grader), under the guidance of **Jerry Ping** (AbbVie) investigated the impact of COVID-19 on Recology waste disposal trends in San Mateo County, a project she first started two years ago. She looked at three collection categories—commercial, residential, and multi-family—and concluded residential waste had the most noticeable increase since the pandemic. Commercial waste has been gradually increasing in the past year after having a significant decrease in March 2020. Multi-family waste disposal was the least affected by the pandemic. Wang also noticed Hillsborough residents had a significantly higher amount of average waste compared to Foster City and San Mateo residents.

One of the reasons the projects were successful is because of the students' eagerness and willingness to contact their mentors and seek guidance.

"It was unlike anything I had ever done at school, as I got to focus on a niche topic that interested me and research it to degrees I never anticipated I would be able to do," said Huang. "I never thought I'd be able to build powerful classifiers, but this project showed me that that was possible."

The experience also provided a valuable and detailed introduction to statistics. Khodakovskaia and Tsvirkunova recall, "During this project, we learned how to analyze data using R-Project software and how to organize data. We additionally covered the importance of statistical testing, how to perform the Chi-squared test, the proportions test, as well as an F-test in linear regression."

Along the way, the students encountered challenges, but they overcame them with confidence with the help of their mentors.

Visit the chapter community site at *https://community.amstat.org/sf-asa/home* for the latest chapter activities. ■

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

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

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

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

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

Also, look for job ads on the ASA website at *https://jobs.amstat.org/jobseekers*.

California

■ The Department of Statistics at UC Riverside invites applications for a tenure-track Assistant Professor position, beginning July 1, 2022. Salary is competitive and commensurate with qualifications and experience. To apply: submit the following to *https://aprecruit.ucr.edu/apply/JPF01457*: Cover Letter, Curriculum Vitae, Statement of Teaching, Statement of Research/Scholarly Activities, Statement of Past and/or Planned Future Contributions to Advancing Diversity, and 3 Letters of Reference. EEO/AA/ADA/Vets Employer.

Florida

■ The Department of Biostatistics and Bioinformatics at Moffitt Cancer Center, a National Cancer Institute-designated Comprehensive Cancer Center in Tampa, Florida, is seeking a tenure-earning faculty member with preference for strong expertise in clinical trial or clinical biostatistics. Methodological research in biostatistics is also expected. To apply, visit our website *MOFFITT.org/Careers* and refer to requisition number #46304. Equal Employment Opportunity Moffitt Cancer Center is an Equal Opportunity/Affirmative Action Employer. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, sexual orientation, gender identity, national origin, age, or protected veteran or disabled status. We seek candidates whose skills, and personal and professional experience, have prepared them to contribute to our commitment to diversity and excellence.

■ The Department of Statistics at Florida State University invites applications for an assistant professor in computational statistics starting August

Institute of Statistical Science, Academia Sinica, Taiwan

Tenure-Track Faculty Positions

The Institute of Statistical Science of Academia Sinica is pleased to invite applications for our tenure-track faculty positions. Academia Sinica, the most preeminent academic research institution in Taiwan, offers a secured research environment facilitated with rich collaboration opportunities as well as the freedom of conducting independent research. With a strong tradition of theoretical and interdisciplinary research, the Institute of Statistical Science is aiming for global excellence in mathematical statistics and various statistical applications.

Applications are invited for tenure-track appointments as Full/Associate/ Assistant Research Fellows (equivalent to Full/Associate/Assistant Professors in Universities) at the Institute of Statistical Science to commence on August 1, 2022 or as soon as possible thereafter. Applicants should possess a Ph.D. degree in Statistics, Biostatistics, Computer Science, Data Science or related areas, and should submit: (1) a cover letter, (2) an up-to-date curriculum vita, (3) a detailed publication list, (4) a research proposal, (5) three letters of recommendation, (6) representative publications and/or technical reports and (7) advisers' names of master and PhD degrees. Additional supporting materials such as transcripts for new Ph.D. degree recipients may also be included. Electronic submissions are encouraged. Applications should be submitted to

Dr. I-Ping Tu
Chair of the Search Committee
Institute of Statistical Science,
Academia Sinica
128 Sec. 2 Academia Road, Taipei
11529, Taiwan, R.O.C.
Fax:+886-2-27886833
E-mail: recruit@stat.sinica.edu.tw

Application materials should be received by **December 31, 2021** for consideration, but early submissions are encouraged.

2022. A doctoral degree from an accredited institution in statistics, biostatistics, or a related field with a demonstrated record of achievement in teaching, academic research and service is required. Please apply at *http://jobs.fsu.edu* (Job ID 50273). The deadline is 12/1/2021. Three letters of recommendation are required. FSU is an Equal Opportunity/Access/Affirmative Action/Pro Disabled & Veteran Employer. FSU's Equal Opportunity Statement can be viewed at: *https://tinyurl.com/ahanbe5f.*

Minnesota

■ The Division of Biostatistics, School of Public Health, University of Minnesota seeks applicants for two tenure-track faculty position at the Assistant Professor rank. We are especially interested in individuals with research interests in causal inference, dynamic treatment regimes, adaptive trial design, risk prediction, biomedical imaging, imaging genetics, and integrative data analysis/data science. Please visit *https://hr.myu.umn.edu/jobs/ext/343694* for additional information or to apply.

■ The Division of Biostatistics, School of Public Health, University of Minnesota seeks applicants for a non-tenure track assistant or associate professor who will provide statistical support for clinical trials coordinated at the Coordinating Centers for Biometric Research (CCBR) and will provide statistical support for a team working on international clinical trials and epidemiological studies. Please visit *https://hr.myu.umn.edu/jobs/ext/343709* for additional information or to apply.

Missouri

■ ATSU is seeking a biostatistician to collaborate with ATSU researchers in recommending study design and data management procedures, performing data analysis, and writing reports for dissemination. A master's degree in biostatistics, statistics, or a related field is required. Residential (onsite), telecommuting, and remote can be considered. Please visit *atsu.edu/atsu-1851* for more information and to apply. EOE.

Pennsylvania

■ The University of Pittsburgh Graduate School of Public Health is accepting applications for Department Chair and Tenured Professor of Biostatistics and Data Science. Additional information is available at *www.join.pitt.edu* (requisition #2100704). Candidates should submit a letter of interest, professional accomplishments, Diversity Equity and Inclusion effort and vision statement, CV and contact information for three referees. Candidates from historically underrepresented communities encouraged to apply.

■ The Department of Biostatistics at the University of Pittsburgh invites applications for a tenure-track assistant professor position beginning fall 2022

The Department of Biostatistics at the Johns Hopkins Bloomberg School of Public Health is seeking an outstanding colleague to join our tenure track faculty at the assistant or associate professor level. We seek candidates to strengthen us in advancing statistical and data science, making discoveries to improve health, and providing an innovative biostatistics education. Responsibilities include methodological and collaborative research, teaching, and mentorship of graduate students. Candidates should also be committed to fostering principles of justice, equity, diversity, and inclusion throughout their research, teaching, and mentoring activities in the statistical and data sciences. Both new PhDs and recent postdoctoral fellows are particularly encouraged to apply.

The Johns Hopkins Department of Biostatistics, founded in 1918, was the first degree-granting department of statistical science in the US and has ranked among the world's best throughout its history. Today, the Department comprises 25 tenure track faculty members, 26 research track faculty, 15 postdoctoral fellows and 79 students, 46 seeking PhDs. Current areas of expertise are diverse, ranging from the real-time analysis of large, streaming data to philosophy and implementation of data science to statistical theory and methods. Our health applications include statistical genomics and genetics, neuroimaging, clinical trials, precision health, environmental health and many others. Learn more at https://publichealth.jhu.edu/departments/biostatistics.

The Department's faculty, students and fellows strive to be influential at the interface of the statistical and health sciences, with the ultimate goal to increase the health of all people. The Johns Hopkins Schools of Public Health, Medicine, and Nursing, the Johns Hopkins Health System, and the Johns Hopkins University are among the top worldwide and provide a research and educational environment in which faculty can achieve scientific excellence. Our faculty have a strong history of being leaders in developing innovative educational content in applied statistics and data science, in particular within the Johns Hopkins Data Science Lab (https://jhudatascience.org). Most importantly, the department prides itself on having a history of strong mentorship for assistant and associate professors, a supportive environment for collaboration and collegiality, and diversity. Margaret Merrell, the School's first female faculty member, was appointed in 1930 and became the School's first female professor. We highly value this tradition: Women and members of under-represented groups are particularly encouraged to apply.

Basic qualifications. Qualified applicants will have a doctoral degree in biostatistics, statistics, computer science, mathematics, computational biology, or a related field. Candidates are required to have their doctoral degree by the time the appointment begins, and the rank of the appointment will be determined in accordance with the successful candidate's experience.

TO APPLY: Submit cover letter, CV, statements on research and educational interests and goals, two manuscripts or articles representing your most important work, and the identity of three references who you have asked to provide supporting letters to https://apply.interfolio.com/97215

In addition, applications should include a statement of demonstrated commitment to the principles of inclusion, diversity, anti-racism, and equity (IDARE) in scholarship, teaching, policy, and practice, and ways to continue to uplift these principles as a member of the Bloomberg faculty.

The Johns Hopkins University is committed to equal opportunity for its faculty, staff, and students. To that end, the university does not discriminate on the basis of sex, gender, marital status, pregnancy, race, color, ethnicity, national origin, age, disability, religion, sexual orientation, gender identity or expression, veteran status or other legally protected characteristic. The university is committed to providing qualified individuals access to all academic and employment programs, benefits, and activities on the basis of demonstrated ability, performance and merit without regard to personal factors that are irrelevant to the program involved. The Johns Hopkins University is a smoke-free environment and as such prohibits smoking in all facilities. The Johns Hopkins University is a drug-free workplace.

or earlier. The primary collaboration responsibility is with the NRG Oncology Statistical and Data Management Center. Apply through requisition #21006614 at *join.pitt.edu* (Direct link: *https://tinyurl.com/uyn8y9ec*). The University of Pittsburgh is an Affirmative Action/Equal Opportunity Employer.

■ The Wharton Statistics and Data Science Department, University of Pennsylvania, seeks a Postdoctoral Researcher. The position is for two years beginning in Summer 2022, with a possible extension to three. The primary focus is for the scholar to develop her or his research. A light teaching load is involved. A PhD is required. Please visit our website to apply: *https://statistics.wharton.upenn.edu/recruiting/dept-postdoc-position*. Direct questions to *stat.postdoc.hire@wharton.upenn.edu.* The University of Pennsylvania is an EOE. Minorities / Women / Individuals with disabilities / Protected Veterans are encouraged to apply.

■ The Wharton Statistics and Data Science Department, University of Pennsylvania, has a full-time, tenure-track assistant professor position, beginning July 2022. Applicants must show outstanding research and teaching skills. Candidates must have a PhD or equivalent (expected completion by June 30, 2023 is acceptable) from an accredited institution. Please apply here: *https://statistics.wharton.upenn.edu/recruiting/facultypositions*. Applying by December 1 is encouraged for full consideration. Forward questions to *statistics.recruit@wharton.upenn.edu.* The University of Pennsylvania is an EOE. Minorities / Women / Individuals with disabilities / Protected Veterans are encouraged to apply.

Texas

■ The University of Texas Health Science Center at Houston School of Public Health seeks to hire two open-rank tenure-track clinical trials faculty. Applicants with demonstrated expertise in leading clinical trials (data and/or clinical components) will be considered. The faculty will also be expected to teach and supervise graduate students. For position details and applicant information, please visit our website: *http://p.rfer.us/UTHBLYB2d.*

CANADA

■ The University of Windsor's Department of Mathematics and Statistics invites applications for a tenure-track position at the rank of assistant professor, in the area of probability and stochastic processes, commencing July 1, 2022. Complete an online application by December 15, 2021, at *www.uwindsor.ca/faculty/recruitment/faculty-postings*. Questions and reference letters can be sent to Dr. Richard J. Caron, Head, Department of Mathematics and Statistics at (519) 253-3000, ext. 3015, Email: *mthsta2@uwindsor.ca.* ■

George Mason University Department of Statistics: Multiple Open Rank Tenure-Track, Tenured, and Term Faculty Positions

The George Mason University Department of Statistics, in the newly formed School of Computing, within the College of Engineering and Computing (CEC), invites applications for multiple tenure-track , tenured, and renewable term (non-tenured) faculty positions beginning Fall 2022. Senior candidates with established records of outstanding research and excellent teaching will be eligible for tenured Associate Professor or Professor positions. George Mason University has a strong institutional commitment to the achievement of excellence and diversity among its faculty and staff, and strongly encourages candidates to apply who will enrich Mason's academic and culturally inclusive environment.

Responsibilities

Successful candidates for tenure-track/tenured positions will be expected to conduct independent and collaborative research; teach and advise undergraduate and graduate students; develop an independent, externally funded research program; participate in all aspects of the department's mission; serve the statistics/data science profession; embrace and help advance the University's strong commitment to diversity, equity and inclusion.

Successful candidates for term faculty positions will be expected to teach undergraduate courses; as well as perform service duties associated with the department's active undergraduate degree programs. Teaching a graduate course or conducting research is optional. Those who bring in research funding may have course reductions. A candidate for Associate or Full Professor is expected to be a leader in curriculum development, undergraduate education, and outreach activities (or conduct research if it is a mixed position with research).

Required Qualifications

Applicants for either tenured, tenure-track or term faculty positions must have received a PhD in Statistics, Biostatistics, or a related field by the start date of the position.

Junior candidates for tenure-track Assistant Professor positions should have demonstrated potential for excellence and productivity in research, and a commitment to high-quality teaching. Senior faculty candidates for tenured Associate/Full Professor positions should have established records of outstanding research and excellence in teaching.

Candidates for term Assistant Professor positions should possess a strong commitment to education and demonstrated excellence in teaching. Applicants at the rank of term Associate/Full Professor must have interests in outreach and significant experience in an academic setting for at least six years post PhD.

Preferred Qualifications

Applicants in all areas of Statistics will be given serious consideration.

For tenure-track and tenured faculty positions, we are particularly interested in candidates with research experience in the areas of data science, data analytics, biostatistics, AI, machine learning, causal inference, computational statistics, networks, and large and unstructured data. Candidates who can build collaborations with other departments within the College, across the University, or national organizations will benefit from the strong support and encouragement for interdisciplinary collaboration from the College and the University. The department and the College place high value on engaging students from traditionally underrepresented groups, and candidates from these groups are especially encouraged to apply.

For term faculty positions, we are particularly interested in candidates with experience in the areas of data science and modern statistics. Preference will be given to candidates with good communication skills and significant teaching experience. Additionally, administrative and/or managerial experience is a plus; as well as research experience. The department and the College place a high value on engaging students from traditionally underrepresented groups, and candidates from these groups are especially encouraged to apply.

About the Department

The Department of Statistics currently houses 22 primary faculty, and has numerous other (adjunct, emeritus, visiting, and affiliated) faculty. It offers a BS degree in Statistics, three MS degrees: Statistical Science, Biostatistics, and Data Analytics Engineering with Concentration in Statistics; and a PhD degree in Statistical Science. The Department of Statistics is renowned with leaders in statistics, biostatistics, and data science/analytics, with credentials in national collaborations, although it is a young department founded in 1992. It is located on the Fairfax, VA, campus in the high-tech corridor of Northern Virginia, 30 minutes from both downtown Washington, D.C., and Dulles International Airport.

The Department of Statistics is involved in multiple projects with INOVA health, one of the leading hospitals in the nation, and with other top corporations. It is also establishing new relationships and research ties via joint initiatives in the greater Washington metropolitan area. These relationships could include research at government agencies on a range of high-impact problems and in several areas of statistics and data science. The department has a distinct potential in becoming a hub for innovations in Statistics and Data Science (SDS) and interdisciplinary research involving medicine, engineering, economics, technology, and society. It is positioned for growth and success. In addition to the positions advertised here, the department is one of the participants of the funded Tech Talent Investment Pipeline, a cluster hire in Computational Systems Biomedicine. Faculty rental housing is also available on campus. Further information about the department is available at http://stat.gmu.edu/.

A Force for Innovation in the Heart of Northern Virginia's Technology Corridor

The College of Engineering and Computing (CEC) at George Mason University is comprised of the Volgenau School of Engineering and a new School of Computing. The College is a fast-growing force for innovation in research and education. Ranked nationally in the top 100 in both undergraduate and graduate education, the College boasts more than 9,100 students in 37 undergraduate, master's, and doctoral degree programs, including several first-in-the-nation offerings. Of the 271 full-time faculty who comprise the College, 91 are tenured, 59 are tenure-track, 89 are instructional faculty, and 32 are research faculty. As part of a nationally ranked research university, its research teams expended $75 million in sponsored research awards in the past year and has projects with over $400 million in current and anticipated awards. The College stands out for its leading research in areas such as artificial intelligence, data analytics engineering, cybersecurity engineering, biomedical imaging and devices, community-based healthcare, autonomous systems, 5G/Next G communications, systems architectures, computational biomedicine, advanced materials and manufacturing, sustainable infrastructure, and more. The College encourages multidisciplinary research and provides ample opportunity for faculty to work with other disciplines.

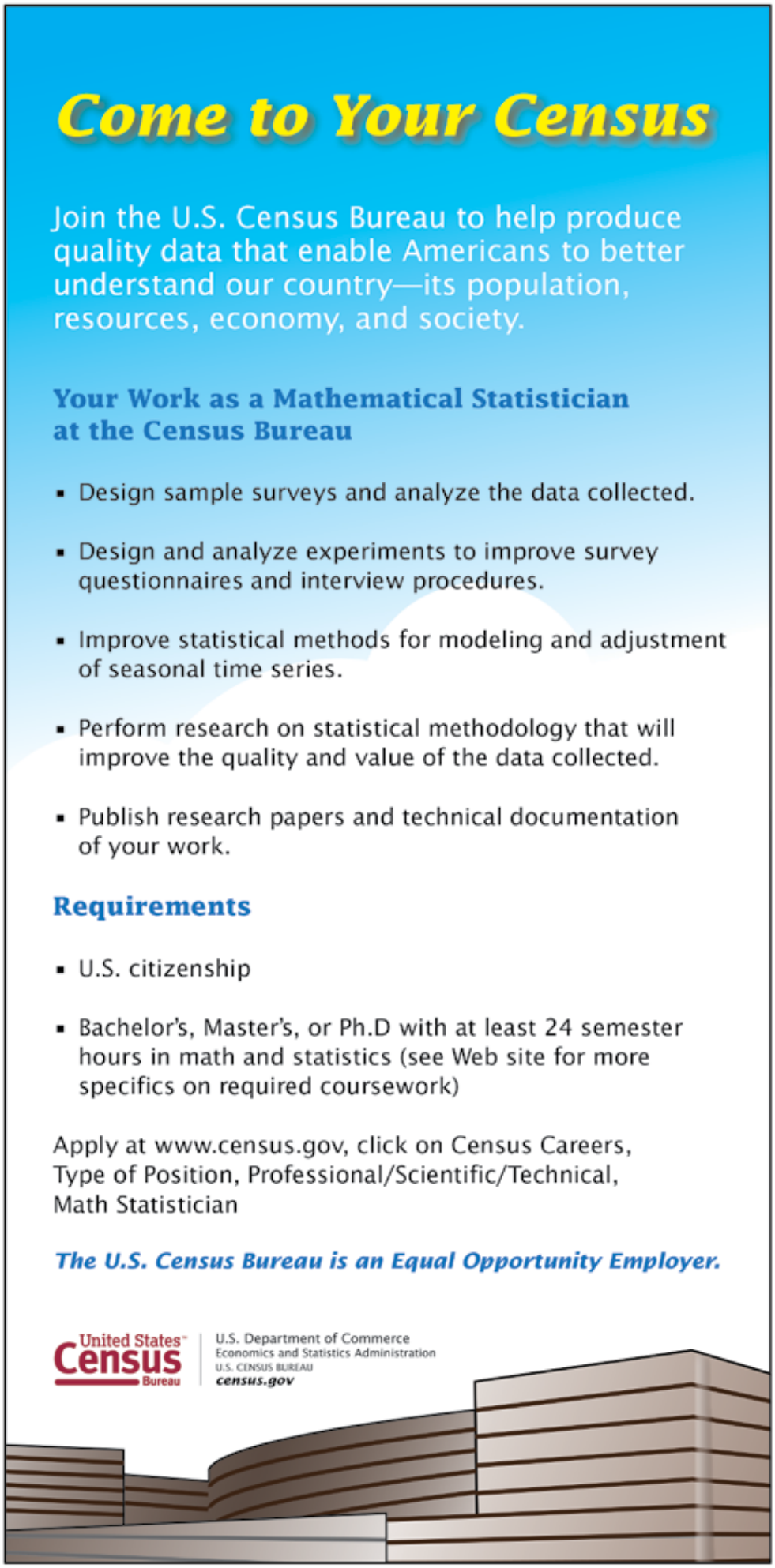
George Mason University is the largest and most diverse public research university in Virginia, with an enrollment of over 39,000 students studying in over 200 degree programs. Mason is an innovative, entrepreneurial institution with national distinction in a range of academic fields. It was classified as an R1 research institution in 2016 by the Carnegie Classifications of Institutes of Higher Education. Mason has campuses in Fairfax, Arlington, and Prince William. Its proximity to Washington, D.C. provides unmatched geographical access to a number of federal agencies and national laboratories. Northern Virginia is also home to one of the largest concentrations of high-tech firms in the nation, providing excellent opportunities for interaction with industry. The region is consistently rated as being among the best places to live in the country, and has an outstanding local public school system.

In conjunction with Amazon's decision to establish a second headquarters in Northern Virginia, the Commonwealth of Virginia announced a multi-year plan to invest in the growth of degree programs in computing. George Mason University has committed to accelerate its plans to grow its capacity in computing and high-tech fields. Among the exciting initiatives being undertaken by the university are the launch of the Institute for Digital InnovAtion, a university think tank and incubator to serve the digital economy, and the expansion of its Arlington Campus with a planned 400,000 square foot building that will house the new Institute for Digital InnovAtion. These initiatives reflect hundreds of millions of dollars in new investment by Mason that will rapidly elevate the university's already leading national position in computing and related areas.

Application

For full consideration, applicants must apply for position number F683AZ for tenure-track/tenured positions, and F685AZ for term faculty positions; at http://jobs.gmu.edu/; complete and submit the online application; and upload a statement of professional goals including your perspective on teaching and research (to attach as 'Other Doc'), a complete CV with publications, a statement on what diversity and inclusion means to you (to attach as 'Other Doc'), and the names of three professional references. The review of tenure-track and tenured faculty applications will begin December 1, 2021, while the review of term faculty applications will begin February 1, 2022. The review of applications will continue until the positions are filled.

George Mason University is an equal opportunity/affirmative action employer, committed to promoting inclusion and equity in its community. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, gender identity, sexual orientation, national origin, disability, or protected veteran status.



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*.

professional opportunities

software

SOCIAL CHATTER

FOLLOW US

community.amstat.org

www.facebook.com/AmstatNews

@AmstatNews

www.instagram.com/AmstatNews

Next month, we'll ask, "If statisticians had a candy bar named after them, what would it be called?" Tag @AmstatNews in your response.

Do you have a creative or interesting question you would like our followers to answer? Share it with us! Email Megan Murphy at *megan@amstat.org.*

What's the best gift to get a statistician?

Eric J. Daza, DrPh (he/him) • @ericjdaza
A train-test set

fatuma mkwepu • @Mkwepu12
Good statistical book for reference

Larry Lesser • @larry_lesser
A gold Markov chain

Gokula Krishnan • @gokstudio
Clearly a version of the Monty hall problem

moria • @moriawong
I thought you were gonna say: amstat membership or ASA giving day

Sharpe Actuary • @Sharpe_Actuary
A Galton board that can do any probability distribution

Dr. Ashley Pitlyk • @ProofOmitted
Anything from Statistical and Mathematical Gifts by NausicaaDistribution on Etsy
www.etsy.com/shop/NausicaaDistribution

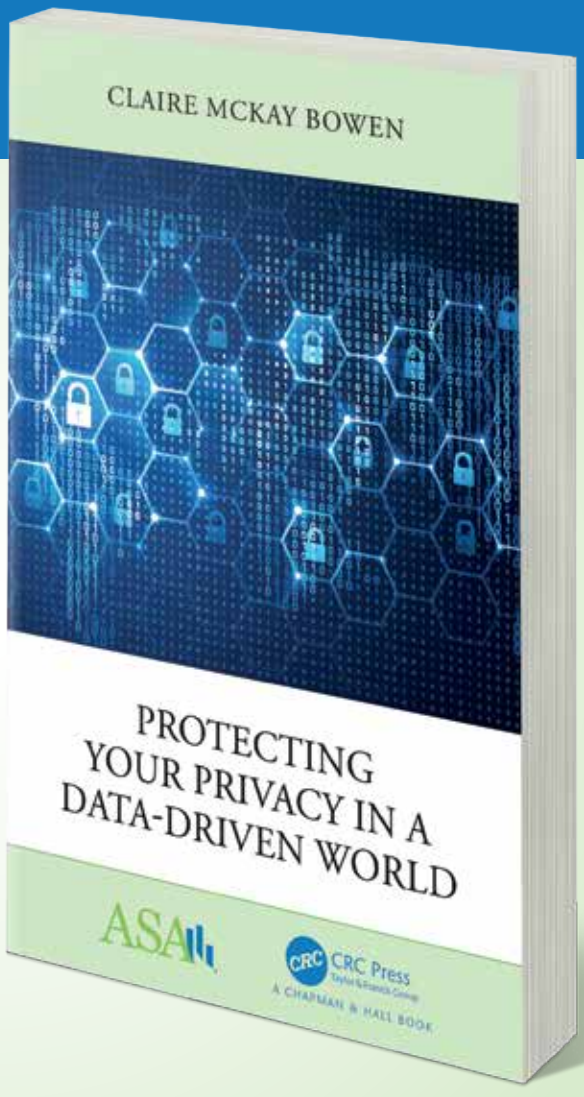
Peter Klaren • @peterklaren
Pint of Guinness

David Mangen • @djmangen
I suspect it varies

PROTECTING YOUR PRIVACY IN A DATA-DRIVEN WORLD

Claire McKay Bowen

SAVE 20%
with code **JML20** on routledge.com



> Staffers and non-technical policy/think-tank researchers need this book. Also data users, like state demographers or public health officials, who need to understand how their data are affected by privacy preserving approaches, but who do not need to know the technical details.
>
> *– Amy O'Hara, Georgetown University*

2022 • 148pp • 9780367640743

CRC Press
Taylor & Francis Group

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.