August 2018 • Issue #494

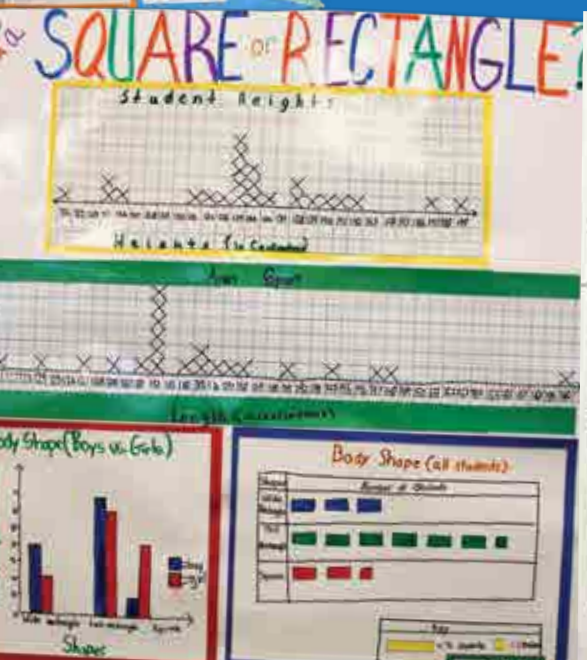
AMSTATNEWS

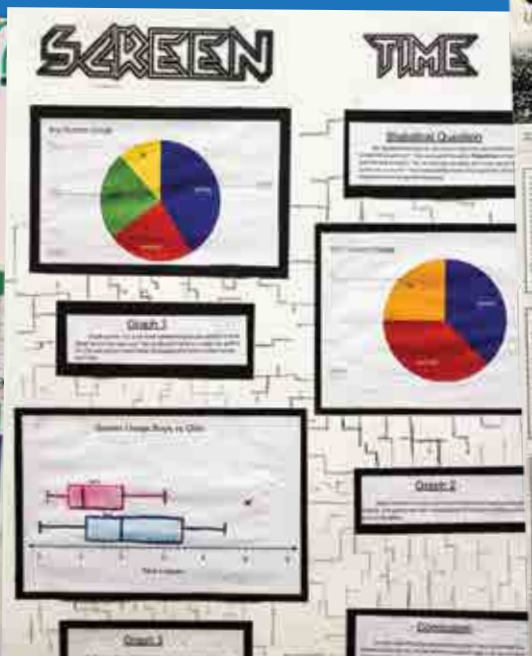
The Membership Magazine of the American Statistical Association • *http://magazine.amstat.org*

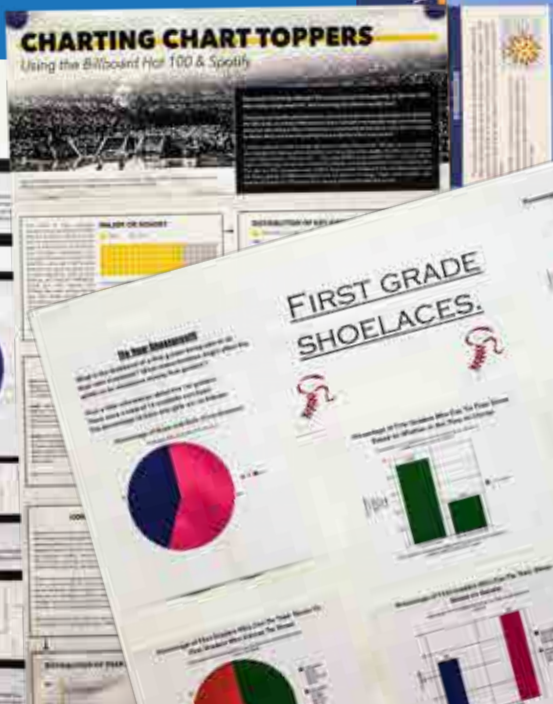
2018 POSTER & PROJECT COMPETITIONS

ALSO: Highlights from 2017 Degree Release: Bachelor's Numbers Close in on Master's

What Does Andrew Althouse Like to Do When He Is Not Being a Statistician?







Stata® and Mata®

The perfect mix of statistics and programming

You know you can do this in Stata

and this

```
. logistic y x1 x2
```

But did you know you can do this?

```
void hello()
{
    printf("Hello, Mata!\n")
}
```

```
x[|1,1 \ 4,4|] = I(4)
```

```
A = (3, 1, 4 \ 1, 5, 9 \ 2, 6, 5)
eigensystem(A, X=., L=.)
```

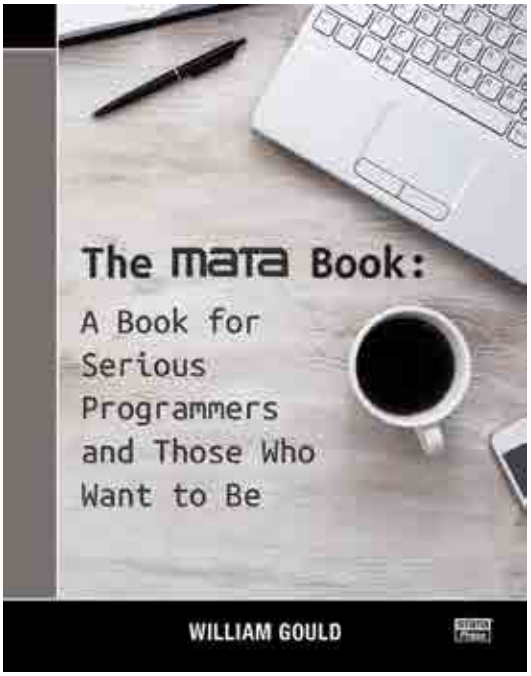
```
p = &a
*p = 4
```

```
url="http://www.somewebsite.net/..."
lines = cat(url)
divs = select(lines, regexm(lines,"<div"))
```

```
struct coord
{
    real scalar x
    real scalar y
}
```

```
real scalar Foo::bar(numeric matrix y)
{
    this.y = y
    this.init()
}
```

```
res = invsym(x)[|1,1 \ 4,4|]
```



The Mata Book: A Book for Serious Programmers and Those Who Want to Be is the book that Stata programmers have been waiting for. Mata is a serious programming language for developing small- and large-scale projects and for adding features to Stata. What makes Mata serious is that it provides structures, classes, and pointers along with matrix capabilities. The book is serious in that it covers those advanced features and teaches them. The reader is assumed to have programming experience, but only some programming experience. As the book says, "being serious is a matter of attitude, not current skill level or knowledge".

Order at **stata.com/amstat-mata-book**.

Learn more at

stata.com/amstat-mata

AMSTATNEWS

AUGUST 2018 • ISSUE #494

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

columns

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

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

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

Modeling Contest Adds Problem, Addresses Energy Use

Starting in 2016, the Consortium for Mathematics and its Applications' annual Mathematical Contest in Modeling added a data insights problem, Problem C. In this new modeling challenge, teams are presented with a modeling problem and data set.

This year's Problem C addressed energy usage of four contiguous US states: California, Arizona, New Mexico, and Texas. Teams were charged with seeking mathematical models that could assist with policy changes for forming a realistic new energy compact focused on increased use of cleaner, renewable energy sources.

New in 2018, the American Statistical Association is designating one outstanding team as the winner of the ASA Data Insights Award.

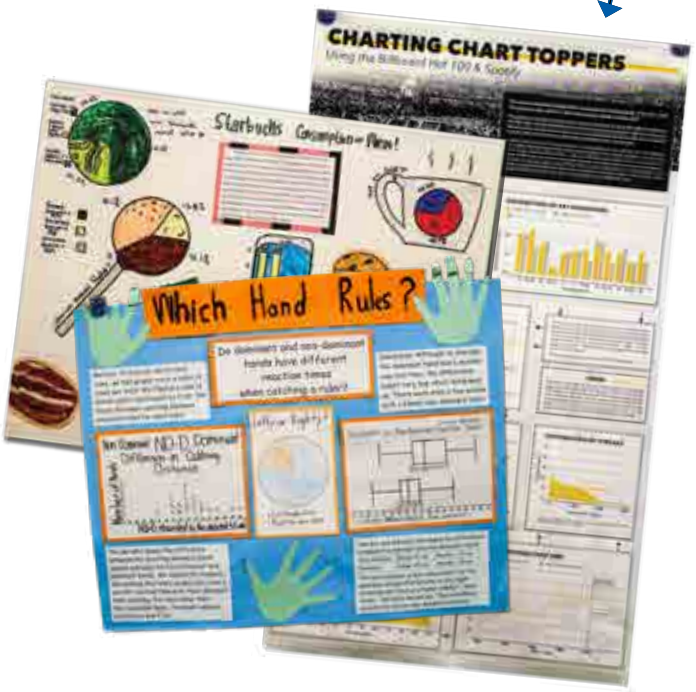
Learn more about the contest and winners at *http://magazine.amstat.org*.

IN MEMORIAM Sadly, **Charles R. Perry Jr.** passed away recently. To read his obituary, visit *http://magazine.amstat.org*.

Make the most of your ASA membership
Visit the ASA Members Only site: *www.amstat.org/membersonly*.

Visit the ASA Calendar of Events, an online database of statistical happenings across the globe. Announcements are accepted from educational and not-for-profit organizations. To view the complete list of statistics meetings and workshops, visit *www.amstat.org/dateline*.

departments



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

ICH E6: Teaching the World to Sing in Perfect Harmony

Statisticians working in clinical trials will be familiar with the acronym GCP, denoting good clinical practice. What may be less well-known is the international regulatory guideline that defines GCP for drug trials, or the organization responsible for issuing that guideline. The International Conference for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) was convened in 1990 with the goal of bringing together pharmaceutical sponsors and regulators from the different regions of the world to develop harmonized guidelines for drug development.

One of the most widely used ICH documents issued early on was ICH E6 Good Clinical Practice. For the past 25 years, this guidance has formed the basis for training individuals of different backgrounds and disciplines working in regulatory clinical trials, especially clinical investigators, study managers, site monitors, and inspectors.

The ASA recognized a few years ago that formal GCP training could benefit the large number of statisticians involved in clinical trials—and not just those working in the pharmaceutical industry or at regulatory agencies. In collaboration with the Drug Information Association (DIA), an ad hoc committee was formed to develop a white paper on statistical training in GCP. Janet McDougall was appointed chair, and Aloka Chakravarty—deputy director of the Center for Drug Evaluation and Research's (CDER) Office of Biostatistics—represented the FDA on the working group. The white paper and subsequent workshop presentations served as a valuable resource when offering preliminary training recommendations for statisticians working in regulatory and research clinical trials. A summary of the committee's work is available at *http://bit.ly/AdHocGCP*.

The timing of this training initiative could not be better, due to changes in the works for E6. But first, a little history (which you know by now I cannot resist giving).

ICH was initially based on the three major regulatory regions of the early 1990s: the US, Europe, and Japan. Representatives from the US Food and Drug Administration (FDA); the European Commission (EC); and Japan's Ministry of Health, Labour, and Welfare (MHLW) and Pharmaceuticals and Medical Devices Agency (PMDA) joined representatives from regulated industry associations—namely the Pharmaceutical Manufacturers Association (PhRMA) in the US, European Federation of Pharmaceutical Industries and Associations (EFPIA), and Japan Pharmaceutical Manufacturers Association (JPMA)—as founding ICH members. In 2015, the organization changed its name to the International Council on Harmonisation and adopted a new management structure (*www.ich.org/home.html*). Membership has expanded over the years to include 16 members and an even larger number of observers. The Chinese regulatory agency, CFDA, joined in 2017, for example, and the International Generic and Biosimilar Medicines Association (IGBA) joined in 2018.

The early products of the ICH were a series of guidelines covering various aspects of trial design and conduct, manufacturing, and quality control. The regulatory agencies adopted the guidelines, which in the US involves an additional step for the FDA to issue formal guidance.

ICH E6 was published in 1996 and describes the responsibilities and expectations of clinical trial stakeholders, including clinical investigators, site monitors, sponsors, institutional review boards, and regulators. It covers the nuts and bolts of running a trial, including data collection, monitoring, reporting, and archiving trial materials. Protecting the safety and rights of patients is addressed through guidelines for the informed consent process and communication of risks (e.g., through investigator brochures).

E6 stood relatively untouched and broadly used for years. A proposal was made to revise it in 2014 to reflect changes that had occurred in clinical trial methodologies and technologies, such as the use of electronic data capture, centralized data monitoring, and risk-based clinical site monitoring. The revision included an integrated addendum and was adopted in 2016. The FDA's corresponding guidance was issued earlier this year.

The ICH meets every six months and rotates meeting locales among the three founding regions.

Ad Hoc Good Clinical Practices Committee Members

Janet McDougall, *Chair*
Scott Berry
Joan Buenconsejo
Aloka Chakravarty
Brenda Crowe
Debra Michaels
Eva Miller
Maria Matilde Sanchez-Kam
Nancy Smith
Terry Weber Sosa

Photo by Jon Gardiner/UNC-Chapel Hill

Lisa LaVange

ICH E8 Working Group in Kobe, Japan

At the June 2016 meeting in Lisbon, Portugal, a group of clinical trialists met with ICH management committee representatives to discuss changes they would like to see in the newly revised E6 document. These suggested changes were outlined in an open letter to the European Medicines Agency (EMA) and ICH that was signed by several European research organizations and an international consortium of health researchers with experience in clinical trials from 19 countries.

As the FDA representative responsible for all ICH documents in the E series (efficacy), I had the opportunity to attend the Lisbon meeting and participate in the E6 discussion with the external stakeholders.

One major concern the stakeholders expressed was the scope of the guidance was too narrowly focused on clinical practice and should more broadly reflect the principles of trial design, conduct, and reporting. We noted during that discussion that E6 does not stand alone as a guide to good *clinical trial* practice, but rather is concerned primarily with practice at the clinical centers of the trial. ICH E9 Statistical Principles for Clinical Trials, for example, provides what could be termed good statistical practice, and many of the topics absent from E6 that were pointed out in the open letter could be found in that guideline.

Other deficiencies noted in the open letter and discussed at the meeting, however, would require a more comprehensive revision to address. In particular, it was thought the current E6 failed to acknowledge that different types of clinical trials posed different levels of risk for participants, and researchers needed to be able to exercise flexibility in managing those risks. Much of the material was pertinent to trials of investigational drugs and did not allow a fit-for-purpose approach for other types of trials. Clinical monitoring in a post-market trial of a marketed drug, for example, may not require the intensity of monitoring or scrutiny of source data that an early phase trial requires, when little is known about potential harm from the drug.

Shortly after the Lisbon meeting, we developed an FDA proposal to renovate E6 and revise the related document, E8 General Considerations for Clinical Trials, in response to the open letter and subsequent discussions with the authors. The proposal was accepted by the ICH, and a reflection paper describing the planned renovation and revision work was posted on the website in January of 2017, inviting public comment (*https://bit.ly/2iUZOCM*). As stated in the reflection paper, "The goal is to provide updated guidance that is both appropriate and flexible enough to address the increasing diversity of clinical trial designs and data sources that are being employed to support regulatory and other health policy decisions. The underlying principles of human subject protection and data quality would remain."

The revision of E8 was proposed to accomplish three goals: (i) introduce a quality by design approach to clinical trials; (ii) expand the scope to nontraditional trial designs and data sources; and (iii) provide a more comprehensive and updated guide to the full list of ICH guidelines pertinent to clinical trials, the complete set of which could be viewed as a reference for *good clinical trial practice*.

Following the E8 revision, the renovation of E6 would begin, further modernizing its approach to GCP. In the reflection paper, we had proposed a series of annexes to E6 corresponding to the use of different data sources (e.g., health care claims data or patient registries, etc.) and trial designs (e.g., pragmatic clinical trials, pharmaco-epidemiology studies, etc.), leaving the main body of the document to cover more traditional randomized clinical trials conducted for product registration.

The revision of E8 began with the ICH meeting in Geneva in June. The expert working group has 30 members from a variety of disciplines. I represent FDA as the rapporteur, and—along with the regulatory chair from the EMA—lead the working group. Mark Levenson, Biometrics Division director in CDER's Office of Biostatistics is FDA's topic lead.

Although not a statistical guideline per se, the importance of statistics in trial design, risk management, data quality, and many other elements of GCP is indisputable. Statisticians have an important role to play as GCP concepts continue to evolve, and the ASA's leadership on the training initiative is timely indeed!

Sincerely,

Lisa LaVange

Staff Spotlight: New Science Policy Fellow Daniel Elchert

He says statisticians must work for the ethical application of statistical concepts

Daniel Elchert is the ASA's new science policy fellow.

ASA Director of Science Policy Steve Pierson along with other ASA executive staff recently chose Daniel Elchert, a PhD student at the University of Iowa, as the ASA's second science policy fellow. Elchert will work to amplify the ASA's science policy efforts to raise the profile of statisticians in policymaking and advocate on behalf of the statistics community.

Originally from Minnesota, Elchert will complete his PhD early this month in counseling psychology. His dissertation research applied hierarchical generalized linear models to analyze whether student behaviors predict college enrollment, dedicating special attention to federal policy implications. He previously earned a master's in educational measurement and statistics, also from the University of Iowa. His undergraduate degree is in psychology from St. Olaf College.

During his graduate training, Elchert has worked as a project manager at the University of Iowa's Center for Evaluation and Assessment, as a doctoral research assistant at ACT, Inc., and as a policy research assistant at the University of Iowa Public Policy Center. For the final year of his PhD studies and after successfully defending his dissertation in May 2017, he is completing an APA-accredited doctoral internship at Michigan State University.

Elchert was selected from numerous candidates for his background of advocacy with policymakers and his interpersonal, communication, and public speaking skills. While working at the University of Iowa Center for Evaluation and Assessment, he advocated for statistically driven program evaluation services to promote state-level educational programing. Elchert also interacted with state-level policymakers while working at the University of Iowa Public Policy Center and further developed his skills in explaining complex statistical concepts to professionals who lacked training in this area.

Elchert believes statisticians must work for the ethical application of statistical concepts in federal policy. He writes, "Due to their training, statisticians are uniquely positioned to describe how concepts like statistical significance and sampling should and should not be applied in policy. This training compels statisticians to authentically communicate such concepts clearly, concisely, and in a manner that highlights the necessity of competent statistical reasoning as part of policy building."

The science policy fellow position presents an excellent opportunity for Elchert to advocate this perspective, collaborate with fellow ASA members, and represent the broader statistics community. ■

Highlights from 2017 Degree Release: Bachelor's Numbers Close in on Master's

Progress Made for Hispanic/Latino Representation in Undergraduate Degrees

The ratio of the number of master's degrees to that of bachelor's closed to 1.2 in 2017, the closest it's been since 1987 when it was also 1.2. The ratio grew to around 2.5 in the mid 2000s.

According to the latest preliminary data release from the National Center for Education Statistics, bachelor's degrees, from 2016 to 2017, grew 22% to 3,398 (36 of which are for biostatistics) and master's degrees increased 4% to 4,059 (693 for biostatistics). Doctoral degrees increased by 5% to 620 (201 for biostatistics), as seen in Figure 1, with the dotted lines showing the associated number of degrees earned by women. Figure 2 shows the comparable data for only biostatistics degrees.

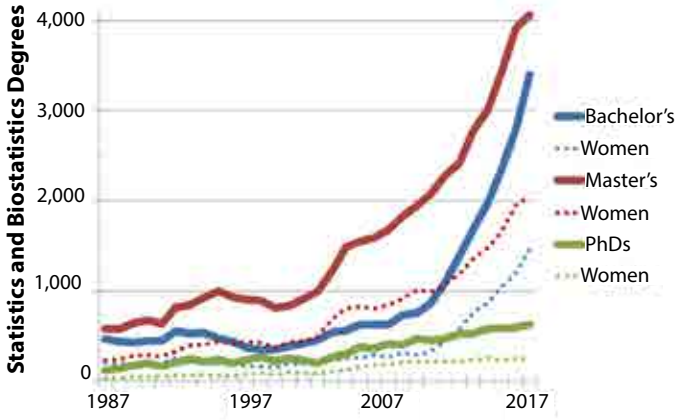


Figure 1
Statistics and biostatistics degrees at the bachelor's, master's, and doctoral levels in the United States. The dotted lines of matching colors are the number of degrees for that degree level earned by women. Data source: NCES IPEDS.

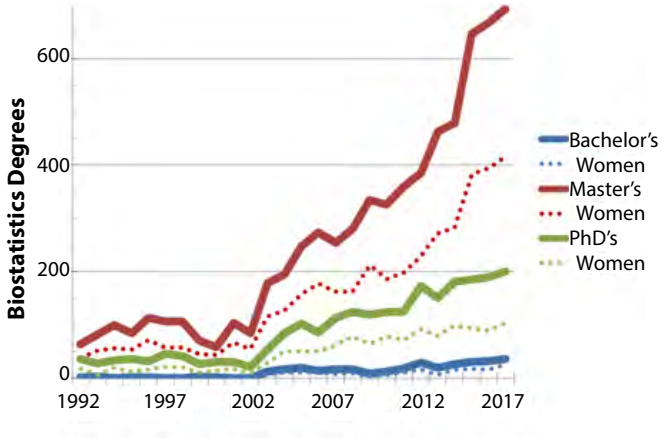


Figure 2
Biostatistics degrees by degree level awarded in the United States. The dotted lines on matching colors are the number of degrees for that degree level earned by women.

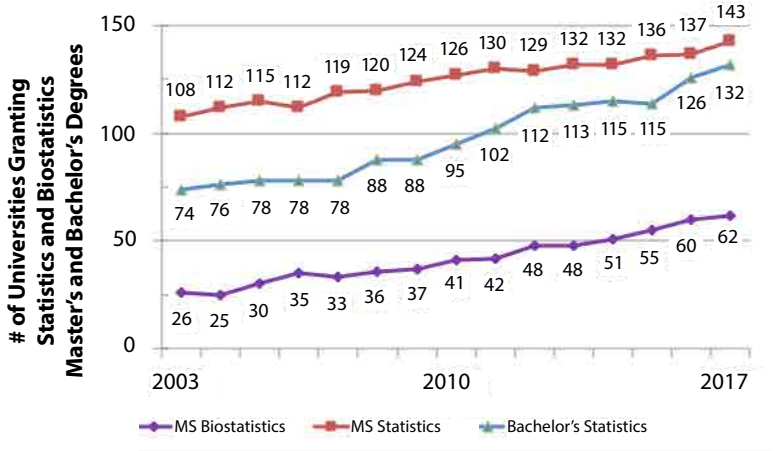


Figure 3
The number of universities granting statistics and biostatistics master's and bachelor's degrees. Compiled from NCES IPEDS data.

Accompanying this growth is an increase in the number of universities granting bachelor's degrees in statistics (from 126 to 132), master's degrees in statistics (137 to 143), doctoral degrees in biostatistics (34 to 39), and doctoral degrees in statistics (69 from a previous high of 67 in 2014), as seen in Figures 3 and 4.

There are 23 universities granting statistics and biostatistics degrees for the first time (at least since 2003) in 2017:

- **Bachelor's degrees in biostatistics** (1): Carnegie Mellon University
- **Bachelor's degrees in statistics** (7): Emory University, Muhlenberg College, Penn State University-Harrisburg, University of Michigan-Dearborn, University of

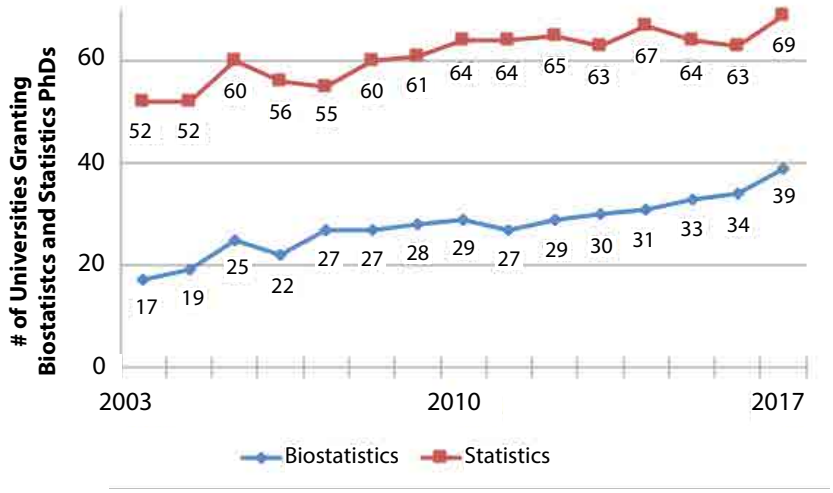


Figure 4
The number of universities granting statistics and biostatistics PhDs. Compiled from NCES IPEDS data.

Table 1

Statistics PhDs	2013	2014	2015	2016	2017	2013–2017	2003–2017
North Carolina State University	12	17	19	20	19	87	258
Iowa State University	15	28	14	9	18	84	182
University of Wisconsin-Madison	13	10	13	24	15	75	191
Stanford University	13	10	4	9	23	59	154
Penn State University	12	9	17	14	6	58	142
Subtotal	65	74	67	76	81	363	927
Total	**379**	**397**	**396**	**402**	**419**	**1,993**	**4,903**

Table 2

Biostatistics PhDs	2013	2014	2015	2016	2017	2013–2017	2003–2017
The University of North Carolina	9	22	24	16	14	85	188
Univ. of Texas Health Science Center	13	7	19	15	13	67	114
University of Michigan-Ann Arbor	15	12	11	11	13	62	153
University of Pittsburgh	12	14	18	5	10	59	140
Harvard University	10	13	12	9	14	58	165
University of Washington	12	9	12	11	14	58	133
Subtotal	71	77	96	67	78	389	893
Total	**151**	**181**	**185**	**190**	**201**	**908**	**2,012**

Washington-Tacoma, Utah Valley University, Valparaiso University

- **Master's degrees in statistics** (7): Central Michigan University, CUNY Queens College, Michigan Technological University, Texas A&M University-Kingsville, University of Houston-Downtown, University of Colorado Denver/Anschutz Medical Campus, University of Kansas
- **Master's degrees in biostatistics** (3): Medical University of South Carolina, New York Medical College, University of Connecticut
- **PhD in biostatistics** (2): University of Miami, Vanderbilt University
- **PhD in statistics** (3): Boston University, University of Texas at Austin, West Virginia University

Table 3

Statistics Master's Degrees	2013	2014	2015	2016	2017	2013–2017	2003–2017
Columbia University	294	287	396	435	441	1,853	3,502
George Washington University	67	155	118	132	117	589	697
Rutgers University	79	94	96	106	78	453	950
University of Illinois at Urbana	61	67	41	75	74	318	630
University of Michigan	55	46	57	74	81	313	665
Subtotal	556	649	708	822	791	3,526	6,444
Total	**2,269**	**2,489**	**2,769**	**3,249**	**3,366**	**14,142**	**29,227**

Table 4

Biostatistics Master's Degrees	2013	2014	2015	2016	2017	2013–2017	2003–2017
Columbia University	37	34	50	68	52	241	442
Boston University	21	45	52	49	49	216	396
University of Michigan	38	24	41	32	40	175	421
Harvard University	24	16	32	36	42	150	244
Emory University	18	17	19	23	36	113	205
Subtotal	138	136	194	208	219	895	1,708
Total	**467**	**495**	**659**	**673**	**693**	**2,987**	**5,864**

Table 5

Statistics Bachelor's	2013	2014	2015	2016	2017	2013–2017	2003–2017
Purdue University	135	197	183	211	199	925	1,191
UC Berkeley	143	160	215	174	215	907	1,453
University of Illinois UC	67	91	111	143	179	591	785
UCLA	50	66	71	127	128	442	541
UC Davis	53	54	60	110	127	404	629
Subtotal	448	568	640	765	848	3,269	4,599
Total	**1,714**	**2,019**	**2,367**	**2,851**	**3,458**	**10,341**	**20,197**

Tables 1–5: Top five universities granting statistics and biostatistics degrees for 2013–2017. These and related data can be accessed at *https://bit.ly/2L5ENUX*.

The top degree-granting institutions over the last five years are in Tables 1-5 for all categories except biostatistics bachelor's degrees (comprehensive list available at *https://bit.ly/2L5ENUX).*

Demographics

Following our practice of alternating demographics updates, we look at the breakdown of degrees for race and ethnicity data and resident aliens and US citizens or residents this year.

As shown in Figure 5, the percentage of master's and doctoral degrees in statistics awarded in recent years to nonresident aliens is approximately 60%. For the same degree levels in biostatistics, it is closer to 50%, while it has grown to about 31% for bachelor's degrees in statistics. The percentages seem generally stable for doctoral-level degrees and on the increase for master's and bachelor's degrees, more so for master's degrees in biostatistics.

Figure 6 shows race and ethnicity data for the degrees granted to US citizens or residents averaged for 2011–2017. (NCES does not report race/ethnicity data for nonresident aliens.) For the five degrees for this subset of the data (except

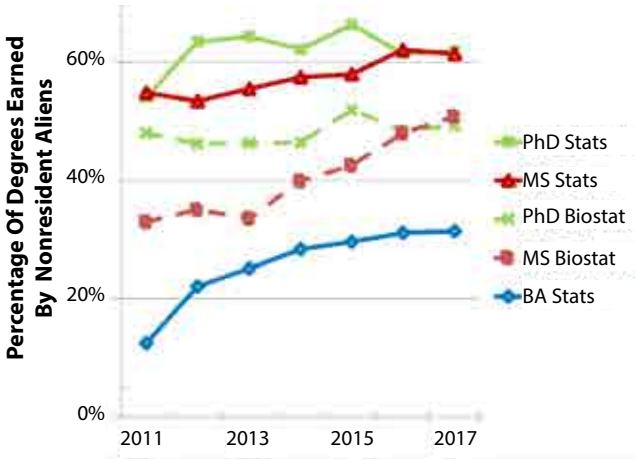


Figure 5
Percentage of statistics and biostatistics degrees earned by nonresident aliens

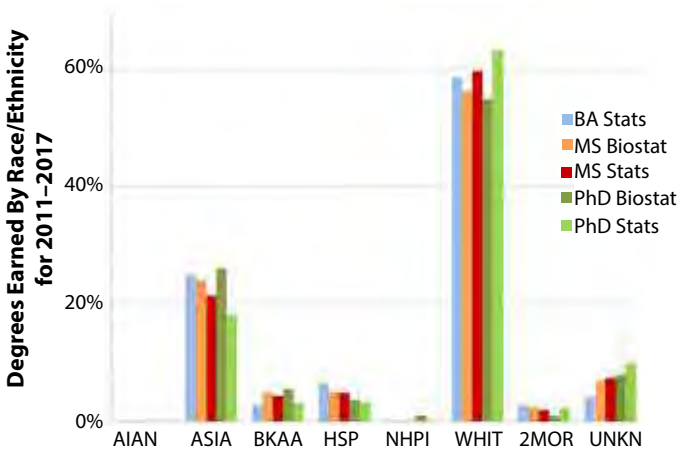


Figure 6
Degrees earned by NCES race/ethnicity group and degree level, averaged over 2011–2017, as a percentage of degrees earned by US citizens or residents

Table 6

	BA Stats	MS Biostats	MS Stats	PhD Biostats	PhD Stats
2011	31 (3.3%)	12 (5.0%)	46 (5.3%)	6 (9.4%)	6 (3.8%)
2012	47 (4.5%)	14 (5.6%)	47 (5.0%)	7 (7.6%)	3 (2.4%)
2013	44 (3.5%)	15 (4.9%)	58 (5.6%)	3 (3.7%)	6 (4.4%)
2014	37 (2.7%)	11 (3.8%)	44 (4.1%)	5 (5.2%)	6 (4.0%)
2015	41 (2.5%)	18 (4.9%)	49 (4.2%)	4 (4.5%)	4 (3.0%)
2016	50 (2.6%)	18 (5.3%)	40 (3.2%)	3 (3.1%)	4 (2.6%)
2017	56 (2.4%)	16 (4.8%)	49 (3.8%)	6 (5.9%)	4 (2.5%)

Number of degrees awarded to African Americans or blacks who are US citizens or permanent residents by degree level. The percentage is the number compared to total number of degrees for that year awarded to US citizens or permanent residents.

biostatistics bachelor's for which the numbers are small), the percentage of degrees earned by those who report their race as American Indian or Alaska Native (AIAN) is essentially 0%. For those identifying as Asian (ASIA), the percentage is around 20%; it is 3–6% for those identifying as black or African American (BKAA). The percentage for individuals of Native Hawaiian or Other Pacific Islander (NHPI) descent is 0–1%. The percentage for those who identify as White (WHIT) is near 60% and 2% for individuals who report two or more races (2MOR). For those identifying ethnicity as Hispanic or Latino (HISP), the percentage is 3–7%. Finally, the percentage for those reporting race/ethnicity unknown (UNKN) is 4–9%.

To better understand the percentages in Figure 6, consider Tables 6 and 7, which show numbers for two under-represented minorities. At the graduate level, for African Americans or blacks who are US citizens or permanent residents, the number of degrees awarded is stable, but the percentage of the degrees awarded to US citizens or permanent residents is generally declining. For the bachelor's level, there seems to be an increase in number, but it's declining as a percentage of overall degrees earned by US citizens or permanent residents.

For Hispanics or Latinos, there again appears not to be any increase in numbers for any degree level except for bachelor's and perhaps master's in statistics. For the former, the increase seems to track—if not surpass—the 145% increase in bachelor's degrees in statistics earned by US citizens or permanent residents from 2011 (943) to 2017 (2,307). In a closer examination of this trend, there were eight institutions that granted more than 10 bachelor's degrees in statistics in 2016 and 2017 to

Table 7

	BA Stats	MS Biostats	MS Stats	PhD Biostats	PhD Stats
2011	44 (4.7%)	7 (2.9%)	33 (3.8%)	3 (4.7%)	3 (1.9%)
2012	43 (4.1%)	19 (7.6%)	46 (4.9%)	1 (1.1%)	7 (5.6%)
2013	64 (5.2%)	12 (3.9%)	59 (5.7%)	3 (3.7%)	4 (3.0%)
2014	86 (6.2%)	14 (4.9%)	50 (4.7%)	2 (2.1%)	2 (1.3%)
2015	130 (8.0%)	18 (4.9%)	57 (4.9%)	5 (5.6%)	3 (2.3%)
2016	151 (8.0%)	22 (6.4%)	60 (4.9%)	2 (2.0%)	6 (3.9%)
2017	157 (6.8%)	17 (5.1%)	69 (5.3%)	7 (6.9%)	8 (5.0%)

Number of degrees awarded to Hispanics or Latinos who are US citizens or permanent residents by degree level. The percentage is the number compared to total number of degrees for that year awarded to US citizens or permanent residents.

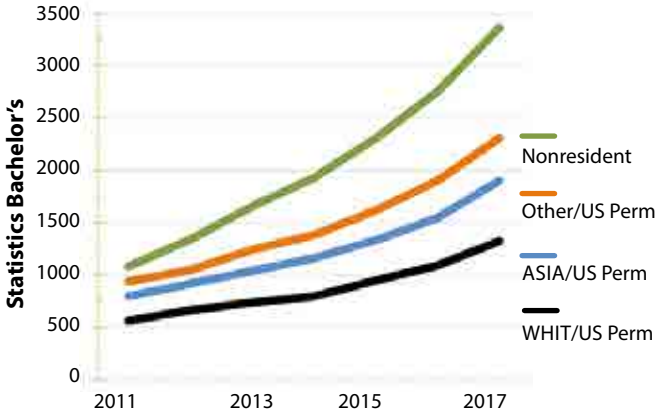


Figure 7
Bachelor's degrees in statistics for the years 2011–2017 by race/ethnicity for US citizen/residents and nonresident aliens. "Other" includes BKAA, HISP, AIAN, NHPI, 2MOR, and NKN.

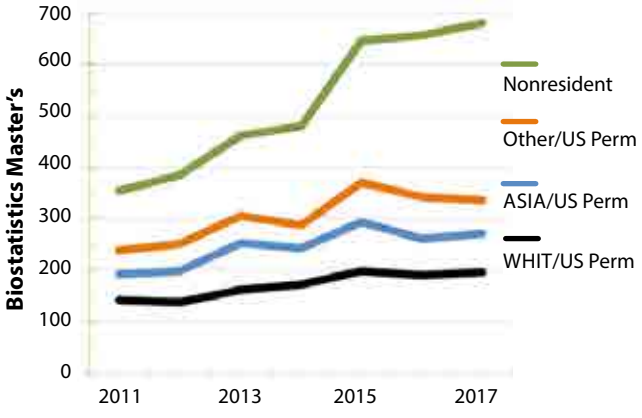


Figure 8
Master's degrees in biostatistics for the years 2011–2017 by race/ethnicity for US citizen/residents and nonresident aliens

students identifying in this ethnic group, amounting to 40% of the total 308 for these two academic years. Seven of the institutions were from California and Florida—University of California (UC) at Santa Barbara, UCLA, University of Florida, San Diego State University, Florida State University, UC-Davis, and UC-Berkeley—with Purdue University being the other institution.

To further understand the growth in statistics and biostatistics degrees, we show the degree growth from 2011–2017 for five degree categories in Figures 7–11. For bachelor's degrees in statistics (Figure 7), the degree growth is substantial for each of the four categories. Within the "Other" category—as shown in Tables 6 and 7—there is strong growth for African Americans or blacks and Hispanics or Latinos. There is also notable growth for two or more races, increasing from 8 in 2011 to 92 in 2017. As another way to represent the growth in numbers for this degree level from 2011–2017, 40% is due nonresident aliens, 33% for white US citizens/residents, 15% for Asian US citizens/residents, 5% for Hispanic/Latino US citizens/residents, 1% for black/African American US citizens/residents, 4% for US citizens/residents of 2 or more races, and 2% for US citizens/residents of unknown race.

For master's degrees in biostatistics, 70% of the growth from 2011–2017 is for nonresident aliens, 17% for white US citizens/residents, 8% for Asian US citizens/residents, and 3% for Hispanic/Latino US citizens/residents (Figure 8).

For master's degrees in statistics, 70% of the growth is for nonresident aliens, 17% for white US citizens/residents, 9% for Asian US citizens/residents, and 2% for Hispanic Latino US citizens/residents (Figure 9).

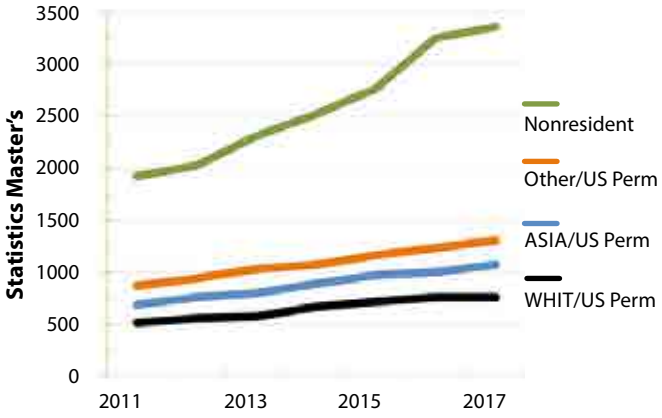


Figure 9
Master's degrees in statistics for the years 2011–2017 by race/ethnicity for US citizen/residents and nonresident aliens

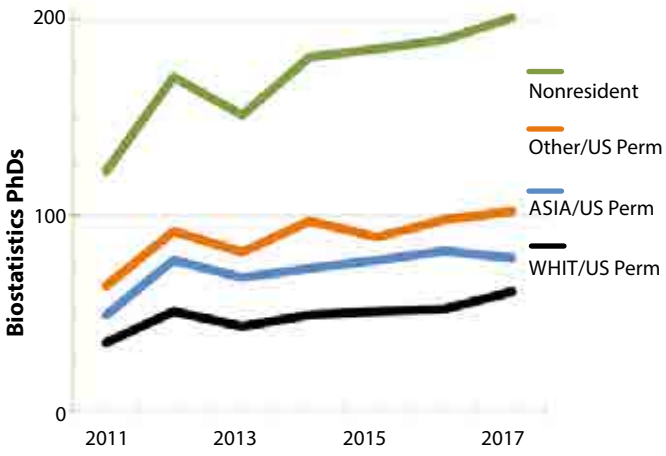


Figure 10
Doctoral degrees in biostatistics for the years 2011–2017 by race/ethnicity for US citizen/residents and nonresident aliens

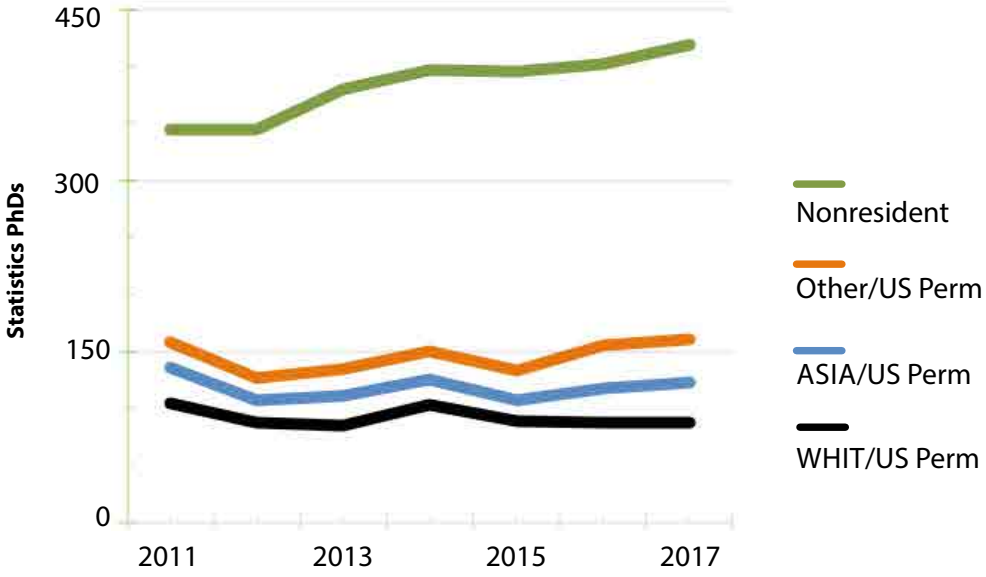


Figure 11
Doctoral degrees in statistics for the years 2011–2017 by race/ethnicity for US citizen/residents and nonresident aliens

For doctoral degrees in biostatistics, 51% of the growth is for nonresident aliens; 33% for white US citizens/residents, 4% for Asian US citizens/residents, 5% for Hispanic/Latino US citizens/residents, 4% for US citizens/residents of 2 or more races, and 4% for US citizens/residents of unknown race (Figure 10).

For doctoral degrees in statistics, 97% of the growth is for nonresident aliens, -23% for white US citizens/residents, 5% for Asian US citizens/residents, 7% for Hispanic/Latino US citizens/residents, -3% for black/African American US citizens/residents, 11% for US citizens/residents of 2 or more races, and 4% for US citizens/residents of unknown race (Figure 11).

For percentage of degrees earned by women in 2017, the level held steady at 43% for bachelor's degrees in statistics, about 49% for master's degrees in statistics, 60% for master's degrees in biostatistics, and 52% for PhDs in biostatistics. For PhDs in statistics, the percentage was down a few points to 34%. For a more complete analysis based on gender, last year's update is available at *https://bit.ly/2LaxU1v.* ■

Turing Award Winner, Longtime ASA Member Publishes *The Book of Why*

Judea Pearl, a longtime ASA member, was interviewed in November of 2012 (see https://bit.ly/2LdNidA*) after receiving the Turing Award from the Association of Computing Machinery. He has recently published a book,* The Book of Why: The New Science of Cause and Effect *(with Dana MacKenzie), that aims to familiarize the general, nontechnical public with recent advances in causal inference. ASA Executive Director Ron Wasserstein interviews him again here to find out what message he thinks his new book sends to* Amstat News *readers.*

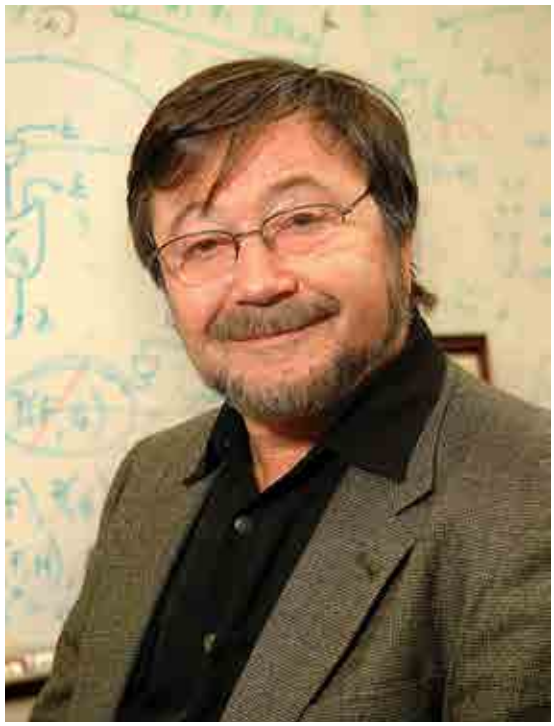
The Book of Why is making a splash in statistics, as well as in machine learning and other data-intensive sciences. I would like to start with a question that you have probably heard many times: What brought you to write the book?

I have official and unofficial answers to this question.

The official answers: First, I have found it both timely and exciting to lay before the public the amazing story of a science that has changed the way we understand scientific claims and yet has remained below the radar to the general public. As we enter the era of big data and machine learning, it is important to share with the public our current understanding of how this new science is likely to affect our lives in the 21st century.

Second, as a part-time philosopher, I have found it intriguing to narrate the history of statistics as viewed from the special lens of its orphaned sister: causation. The story of this "forbidden love" was never told before and, believe me, it is full of mystery, intrigue, personalities, dogmatic orthodoxy, and heroic champions of truth and conviction.

Finally, my unofficial reason is to incite a rebellious spirit among rank-and-file statisticians, so the excitement that currently fuels causality research in academia percolates down to education and to practice. In other words, I am impatient with the slow pace at which the tools of causal inference are becoming an organic part of statistical thinking.



Judea Pearl's *The Book of Why* is a nontechnical book for the general public that discusses recent advances in causal inference.

You expressed a similar impatience in our interview six years ago. And you have initiated the ASA Causality in Statistical Education Award to close the growing gap between research and education. Hasn't this initiative met your expectations?

It has. But, with age, my impatience grew stronger and less forgiving. Of course, the availability of instructional material made it easier for instructors to introduce aspects of causal inference in graduate

courses, but it was not sufficient to change the curriculum of undergraduate classes. Nor was it sufficient to reshape the minds of practicing statisticians or high-profile academics who are too busy to sort out what all the causal inference "hype" is about.

What *The Book of Why* is doing can be described as "the democratization of causal inference." It awakens the untrained students to the realization that "it's easy and who needs the 'experts' and all their quibbles?" As a result, the book is accomplishing what I have failed to achieve in the past 30 years through hard labor and scholarly discussion with the leading statisticians of our time—a mass uprising of common sense.

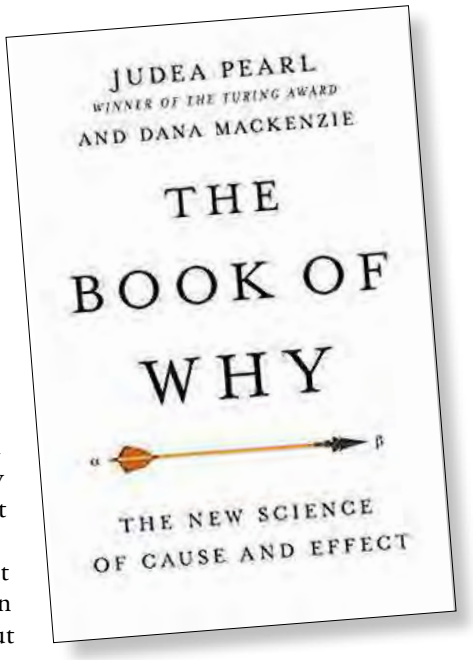
I have read that some statisticians find your claims to be "hard to swallow," especially your characterization of causal inference as "The Causal Revolution" and your depiction of statisticians as antagonistic to causal thinking. Can you comment on these sentiments?

These are not only sentiments but natural complaints voiced by practicing statisticians who are genuinely surprised by how the history of statistics is viewed from the causal lens.

Take for instance the mantra "correlation does not imply causation," which every statistics student has learned to chant, demonstrate, and internalize. *The Book of Why* dissects this mantra to far-reaching conclusions that seem indeed "hard to swallow," even to seasoned statisticians.

First, it can be strengthened to assert that no causal conclusion can ever be obtained without some causal assumptions (or experiments) to support the conclusion. This is hard to swallow because it sounds circular, and because if you look at the statistical literature from 1832 to 1974, you will find many ideas about what is needed to substantiate causal conclusions (e.g., Yule, Fisher, Neyman, Hill, Cox, Cochran), but not one causal assumption—at least not formally.

This raises an interesting question: Why could not these giants of statistics come up with a simple principle, telling us what assumptions are needed for establishing a given conclusion, and let us judge—for any given situation—whether it is plausible to make those assumptions? And here comes the second surprise that is even harder for people to swallow: Even if they knew the needed assumptions, statisticians could not have articulated them mathematically—they simply did not have the language to do so.



Readers refuse to accept this linguistic deficiency until I ask them to write down a mathematical expression for the sentence, "The rooster crow does not cause the sun to rise." Failing this elementary exercise drives people to realize a totally new notational system is needed; the beautiful and powerful language of probability theory and its many extensions cannot make up for this deficiency.

The needed notation first came into being in 1920, when the geneticist Sewall Wright put down on paper a new mathematical object: a causal diagram. Thus, statistics was separated from causality, not by antagonism or disdain, but by a language barrier—the toughest barrier for humans to acknowledge and to cross. Now that the barrier is behind us, it is only natural we should call the crossing a "Causal Revolution."

These are interesting theoretical points, but I wonder if they are likely to have significant impacts on the practice of statistics or on statistical education.

The most significant practical impact of the Causal Revolution would probably be a continuous erosion of the supremacy of randomized clinical trials (RCT) in the development and evaluation of drugs, therapeutical procedures, and social and educational policies. Last year, for example, the editors of one of the two leading medical journals in America stated that authors should not talk about causation unless they have conducted a randomized clinical trial.

Miguel Hernan of Harvard and several other specialists in public health vigorously protested this restriction, and Hernan wrote, "The biggest disservice of statistics to science has been to make 'causal' into a dirty word, the C-word that researchers have learned to avoid."

Indeed, considering the practical difficulties of conducting an ideal RCT and its inherent sensitivity to sample selection bias, observational studies have a definite advantage: They interrogate the target populations at their natural habitats, not in artificial environments choreographed by experimental protocols.

The book is accomplishing what I have failed to achieve in the past 30 years through hard labor and scholarly discussion with the leading statisticians of our time—a mass uprising of common sense.

The development of a new toolkit that allows scientists to estimate causal effects from observational studies now opens a wide variety of applications—from medicine to social science to ecology—free from problems of ethics, costs, and external validity that plague randomized clinical trials.

True, observational studies are necessarily sensitive to modeling assumptions that must be defended on scientific grounds. However, the transparency with which those conceptual assumptions are displayed, coupled with the ability of testing them against data, now make observational studies serious contenders to RCTs.

I would like to go back to education and ask what you believe would induce a typical statistics instructor to introduce aspects of causal inference in a standard statistics class.

Curious students who read *The Book of Why* will make it impossible for statistics instructors to skip such aspects.

Take for instance Simpson's paradox, a phenomenon discussed in every statistics class, usually for the purpose of demonstrating that "correlation is not causation." The discussion usually ends with a song of praise to statistical tables for showing us that the reversal can indeed occur in the data, hence the paradox does not exist. Done. Some instructors go a bit further and praise the table for protecting us from naïve beliefs in miracle drugs that are good for men, good for women, and bad for the population.

Now imagine an inquisitive student raising his/her hand and asking the very obvious question: So, what do we do if we find Simpson's reversal in the data? Shall we believe the aggregated data or the disaggregated data? I do not believe any instructor would in good faith be able to evade this question, suspecting the student knows the answer; it takes a few lines to describe. In other words, instructors would not be able to skip the causal implications of Simpson's paradox, as their professors did to them.

The same applies to Lord's paradox, spurious correlations, instrumental variables, confounders, and other causal concepts that were used to embarrass statistics instructors in the past.

The graphical approach you advocate in the book is but one of several approaches currently used in causal inference. Would a reader versed in potential outcome analysis feel comfortable with your methodology?

Not only comfortable, but enlightened and liberated. Researchers entrenched in potential outcome analysis will discover, to their amazement, that the following three notorious weaknesses of potential outcomes can easily be overcome:

- Assumptions of "conditional ignorability," which currently underlie every potential outcome study, can be made not because they facilitate available statistical routines, but when they are truly believed to hold in the world. They are, in fact, vividly displayed in our model of the world (i.e., the causal diagram), where they can be scrutinized for plausibility, completeness, and consistency.
- When assumptions of "conditional ignorability" do not hold, it is not the end of the world; the analysis can continue, and causal questions answered using other types of assumptions the model may license.
- Modeling assumptions need not remain opaque or data-blind; they can be tested for compatibility with the available data, and the model tells us how.

Making these three bullets available to researchers from the potential outcome camp will break through a wall of cultural isolation and enable them to communicate with the rest of the research community in a common, unified language.

To summarize, the democratization of causal inference is bringing about a globalization of common sense and a breakdown of cultural barriers. I am gratified to see *The Book of Why* contributing to this process. ■

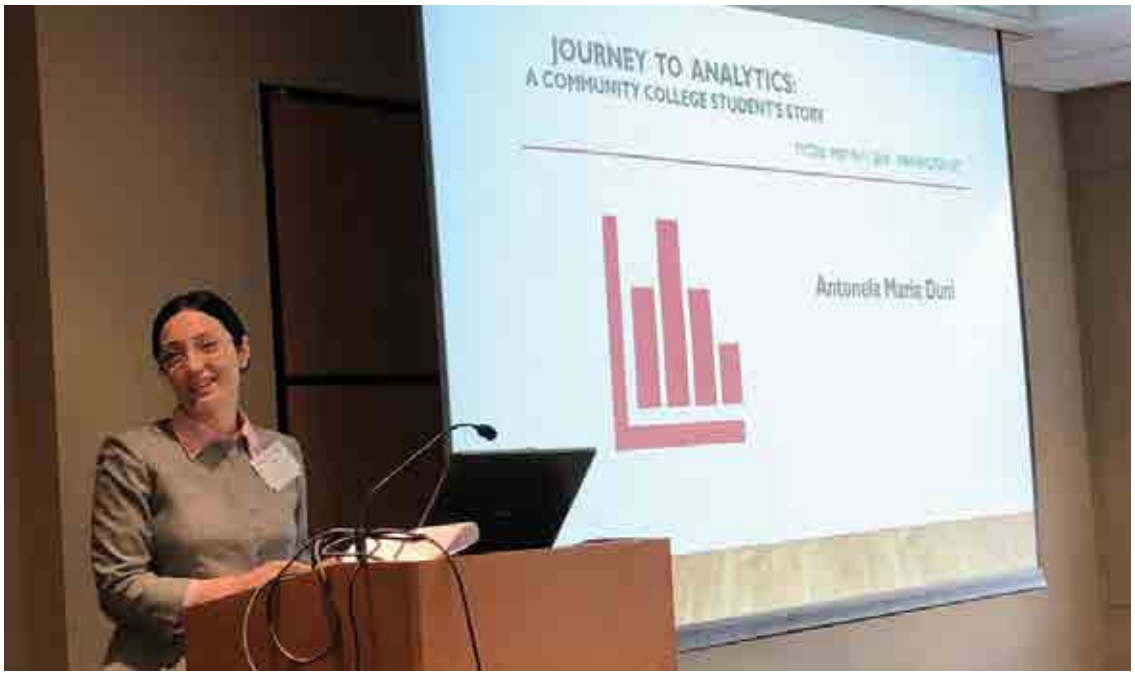


Photo by Nicholas Horton

Two-year college data science student Antonela Duni provides her perspective at the summit.

ASA Hosts Summit to Discuss Data Science Curricula for Two-Year Colleges

Seventy-three representatives from two-year colleges, four-year colleges and universities, government, and industry gathered in May to discuss the future of data science education at two-year colleges.

Funded by the National Science Foundation in an award to the American Statistical Association, the Two-Year College Data Science Summit (TYCDSS) took place in the facilities of Booz Allen Hamilton (BAH) in Arlington, Virginia, May 9–10. The summit, convened by the ASA, included representatives from professional associations such as the American Mathematical Association of Two-Year Colleges (AMATYC), Association for Computing Machinery (ACM), and Mathematical Association of America (MAA). The summit also included those who have been working to develop educational pathways in data science, including the Park City Math Institute Undergraduate curriculum guidelines (see *https://bit.ly/2lqW7rT*), National Academy of Sciences undergraduate report on data science, and Oceans of Data Institute's profiles of big data and data science practitioners.

The primary goal of the summit was to develop sample curricula that could be used to guide the development of future programs. Discussions focused on the key principles and practices that would be appropriate for the following three pathways/themes:

1. Two-year degrees intended for transfer to four-year colleges and universities
2. Two-year degree programs intended for students who wish to go directly into the workforce
3. Certificate programs intended for professional development

The secondary goals of the summit were to summarize the current state of data science/analytics programs at two-year colleges, identify challenges in establishing such programs at two-year colleges, and identify necessary resources for two-year colleges when creating a data science/analytics program.

The goals of the summit were met through active discussion and engagement. Participants were split into six discussion groups, with two groups devoted to each of the themes. Each discussion group had representatives from the communities above, when possible. Groups were charged with generating program outcomes and defining the intended student audience. Time was given for inter-group discussions and feedback.

Rob Gould

Roxy Peck

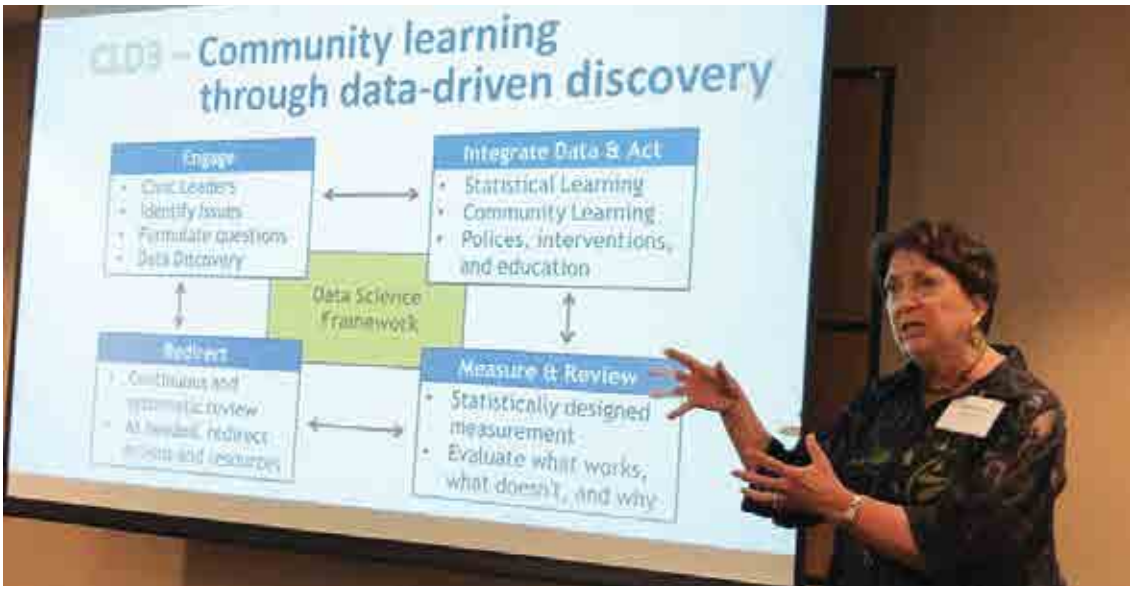


Photo by Nicholas Horton

Sallie Keller, who started her undergraduate career at a two-year college, delivers the keynote for the summit.

Steering Committee Members

Rob Gould, University of California at Los Angeles (program co-chair)

Roxy Peck, California Polytechnic State University, San Luis Obispo (retired) (program co-chair)

Beth Hawthorne, past-chair of the ACM Committee for Computing Education in Community Colleges and vice chair of the ACM Education Board

Nicholas Horton, Amherst College

Randy Kochevar, Oceans of Data Institute

Brian Kotz, Montgomery College and chair of the AMATYC Data Science Subcommittee

Mary Rudis, Great Bay Community College

Brad Thompson, Delaware Technical Community College

Heikki Topi, Bentley University

ASA staff support included **Donna LaLonde** and **Steve Pierson,** who served as the NSF award principal investigator.

A number of key themes emerged from the process. All agreed there is need for a workforce capable of analyzing, storing, managing, and sharing data. The consensus of the participants was that two-year colleges will play an important role in filling this need. Several participants spoke to the demand of hiring great numbers of data professionals at the associate level in the near future, and, in particular, thought this need required a system of strong data science programs at two-year colleges.

After welcoming comments from program co-chairs Rob Gould of the University of California at Los Angeles and Roxy Peck—retired from California Polytechnic State University, San Luis Obispo—as well as NSF program officer Corby Hovis, Sallie Keller—founder and head of the Virginia Tech Social and Decision Analytics Laboratory and 2006 ASA president—gave a kick-off plenary talk. She began her undergraduate education at a two-year college and transferred to complete a bachelor's degree before subsequent graduate training. Keller described the data science work her laboratory does and how it demonstrates the societal importance of statistics and data science.

Antonela Duni, a student and spring 2018 graduate of Great Bay Community College—one of the first colleges to offer a certificate in data science—provided the student perspective. Aaron Burciaga, vice president of data science and artificial intelligence at Booz Allen Hamilton, also spoke to the group, sharing his enthusiasm for two-year colleges providing data science training. Additionally, fellow BAH employee and prominent data scientist Kirk Borne participated in the entire summit.

Writing groups are formulating reports of the three themes, with drafts expected this summer. The organizers have also set up a Google Group for participants and others to share resources and work to build a broader network of educators.

Participants were enthusiastic about the summit, for its usefulness—both in content and networking—and organization. Many urged it be an annual workshop to foster the growth of the emerging community interested in two-year college data science education.

For more about TYCDSS, visit *https://bit.ly/2uePt9Y.* ■

Census-FNS-ERS Partnership Delivers Unique Products by Leveraging Data

Mark Prell and Erik Scherpf

Two statistical agencies—the US Census Bureau and US Department of Agriculture's (USDA's) Economic Research Service (ERS)—have formed a strategic partnership with USDA's Food and Nutrition Service (FNS), which administers USDA's food and nutrition assistance programs. The Census-FNS-ERS Joint Project is a long-term effort to acquire state-level administrative data for USDA nutrition assistance programs and to link those data to Census Bureau surveys and other administrative files. The project leverages the strengths of multiple data sources, obtaining statistical results that could not be achieved using any one source by itself.

Results of the project include specialized statistical tables and data visualizations that help states administer USDA nutrition assistance programs, as well as findings on policy-relevant issues concerning program access, targeting of benefits, and participation patterns.

The Census-FNS-ERS joint project acquires data from two of USDA's largest food assistance programs—the Supplemental Nutrition Assistance Program (SNAP, formerly known as Food Stamps) and the Special Supplemental Nutrition Assistance Program for Women, Infants, and Children (WIC). Last year, SNAP served about 42 million low-income people per month at an annual cost of $68 billion to taxpayers. WIC serves low-income women, infants, and children at nutritional risk. In 2016, WIC served an estimated 48% of all infants born in the US. The program's annual cost is about $5.6 billion.

The joint project is a national undertaking. State SNAP and WIC agencies collect and store the administrative data for the programs. State agencies can participate by sending their confidential microdata (involving millions of records that must be maintained in a secure data environment) to the Census Bureau in exchange for state-specific tables and data visualizations. The project began in 2012 with three states and has grown to 23 SNAP agencies and 11 WIC agencies.

Any data source has strengths and limitations. The states' SNAP and WIC data are known for their accuracy. However, these administrative records exclude people who do not participate in the program. The Census Bureau's survey data have a complementary strength because they include both program participants and nonparticipants. Survey data also include a rich set of socioeconomic variables, such as annual income—critical for certain analyses. Survey-based data on program outcomes, such as health and food security, further expand research possibilities. Matching these survey and administrative data enhances the ability to conduct research previously impossible.

The project is working to scale up by acquiring and linking data from more states, agencies, and programs. The Census Bureau acquires state administrative data on Temporary Assistance to Needy Families (TANF). The bureau has also developed data-sharing agreements with the US Department of Veterans Affairs (VA), US Department of Housing and Urban Development (HUD), and Medicaid and Medicare programs in the US Department of Health and Human Services (HHS). Data on veterans, housing, and health and human services programs can be linked to state administrative records for USDA nutrition assistance programs, to surveys, and to the Decennial Census. In addition to acquiring more data, the joint project is making data more available to academic researchers through the Federal Statistical Research Data Centers (FSRDCs).

States that participate in the project receive specialized tabulations designed to address a key policy-relevant issue: How effectively does SNAP (or WIC) reach the people the program is designed to serve? Using the linked data and a methodology developed in an ERS report (*https://bit.ly/2KRk73K*), the bureau estimates the SNAP access rate—the percentage of people who receive program benefits (according to administrative data) among those estimated to be eligible for SNAP (based on survey data); an analogous WIC access rate is provided to state WIC agencies. SNAP and WIC access rates are estimated at the state level, at highly granular geographic levels within a state, and for demographic subgroups.

Prior to this project, it was not possible for states to obtain these results. A related product, which is newly developed, is a SNAP Profile—a web-based, interactive data visualization tool that provides access rates and other SNAP information; a SNAP Profile for New York State provides a first case (*https://bit.ly/2LdOZaP*).

Using the linked data, ERS researchers address policy issues such as program access, targeting of benefits, and participation patterns. For example, using linked data and careful modeling of membership in SNAP households, an ERS report (*https://bit.ly/2JvVbcG*) found that SNAP benefits are targeted more toward those in deep poverty (with annual income less than 50% of the poverty line) than is detected using survey data alone.

The Census-FNS-ERS project exemplifies how a partnership between statistical and program agencies, can succeed by serving the mission of each partner. The project brings together data, resources, and expertise to create unique statistical products that are valued by policymakers and other stakeholders in the programs. ■

Mark Prell is a senior economist at the Economic Research Service (ERS), US Department of Agriculture (USDA), where he coordinates interagency activities that support the Census-FNS-ERS Joint Project.

Erik Scherpf is an economist at the Economic Research Service (ERS), US Department of Agriculture (USDA), where he leads and conducts statistical research using the project's administrative and linked data.

Editor's Note
The opinions expressed here are those of the authors and should not be attributed to the US Department of Agriculture, US Census Bureau, or federal government.

NCHS Director Addresses Challenges for Federal Statistical Agencies

Charles Rothwell is director of the National Center for Health Statistics. He is a Fellow of the ASA and recipient of the Hal Dunn Award in Biostatistics. He also served as a Captain in the Marine Corps. Rothwell earned his BS in physics from the Virginia Military Institute, his MS in operations research and systems analysis from The University of North Carolina, and an MBA from the University of Maryland.

Editor's Note
These are the remarks Rothwell made to members of the Council of Professional Associations on Federal Statistics.

It is now very close to four and a half decades since I first became involved in health statistics, which at that time was in the state health department in North Carolina. It was a watershed time for the health care sector and especially for public health. Finally, IT—we called it data processing back then—started to have a significant impact on our ability to monitor health status. Although a watershed time, it was a gradual tide, which rode on the crest of new technologies. I was fortunate enough to be carried along with that tide.

We are faced with another watershed time … now for federal statistical agencies. Bob Groves—the provost at Georgetown, former director of the US Census Bureau, and perhaps one of our country's finest survey statisticians—used this very term at a recent National Academies presentation I attended and it was not hyperbole.

At a time when unbiased data are needed more than ever to document where we are as a nation—what has worked and what has not—and to help guide where we need to go as a nation, societal and technological changes are putting the way we collect and provide these data back to the public, decision-makers, and researchers at risk.

Much of the problem rests with the public who volunteers to provide us this information. They have reservations about government, they do not know us, they are being bombarded by requests for information and they see failures of trusted businesses releasing their personal data inappropriately.

Perhaps of equal importance is that the way they consume data and information has drastically changed in just the last five years. Think of your bookstores … how about your local library … major newspapers losing subscribers … how people are accessing information has changed, and those choices affect how we as federal statistical agencies fit in. If we are not giving the public the information they need in the way they now access and consume information and if they do not know the data come from us because of our lack of branding … why should they commit time to us for our surveys?

We must overcome these changes by providing them with information they can use through their personal assistants, apps, bots … We need to meet them where they are, or we as a nation—as a society—won't have the data to govern appropriately. Our anonymity is no longer a virtue. As in voting, we need to advertise to the public why participation is important and the difference participation can make in the lives of their family members, friends, and communities.

At NCHS, all our statistical systems need to be changed—significantly changed over the next five to six years. The problem we face is that our statistical systems are needed in the field while these major changes take place. Why? They are the gold standard other data collection systems—both government and private—adjust themselves to. That means we need funds to test, to innovate, while we continue to provide needed data now as best we can. Without these funds to innovate in a significant way, all we can do is innovate around the edges. We can keep our trains going down the track, but that track is leading us ever so surely to a dead end. We should be experimenting with flying!

So, let me talk about the challenges for NCHS, which are quite similar to those faced in other statistical agencies:

1. **No funding for significant innovation for statistical agencies.** Without funding, we will need to close down major surveys to make that change. We were lucky at NCHS when funding through PCORI (Patient-Centered Outcomes Research Institute) allowed us to bring a significant change to the timeliness of vital statistics. That funding will soon end.

2. **Do we always need to say something nationally before we publish and make data available? How good is good enough?** In vital statistics, we are now reporting drug overdose deaths on a monthly basis for the US and a few states with good reporting on specific drugs used. Such incomplete and provisional reporting is highly useful for public health surveillance and encourages faster and more complete reporting by other states. If such incomplete reporting has such value, why not elsewhere? Electronic health records are our future in the reporting of health care in the US. Instead of tens of thousands of records through surveys, we could conservatively have tens of millions of records. There are, though, significant

issues of data compatibility across vendors of these systems. Why not work with a single vendor or with a large health care conglomerate and report on findings just with their data? Such incomplete reporting could still have significant use on the treatment of rare conditions and, as with vitals, encourage other vendors to provide standardized data that will lead to national reporting on detail never before achieved.

3. **Computing platforms for large data sets: Whether it means analyzing millions upon millions of electronic health records or genomic data or data derived from wearables, we need larger and more flexible computing platforms than our current fixed server farms.** Whether it is using a highly secure cloud environment or a large government-controlled computing environment such as in Oak Ridge, we need the capability to quickly ramp up with computing power and appropriate analytic software.
4. **New staff with new skills.** We need the type of staff who can take advantage of the new computing platforms while we have existing staff maintain existing systems. We especially need staff interested in using data visualization software to make more effective our quality control systems, as well as with data dissemination.
5. **Response rates.** Response rates continue to decline for in-person surveys. We need more research into management of nonresponse bias. We need to be even more creative in the use of paradata to improve response rates. We need to look at other survey modalities to augment or supplement our surveys. We need to become known through advertising our brand and providing our data back as information to the public for their mobile devices and automated personal assistants.
6. **Rethink our surveys**. How might NHANES (our physical measurement survey) be thought of as part of NHIS (our household interview survey)? What physical measurements could now be undertaken in the NHIS? Could NHANES (National Health and Nutrition Examination Survey) measurements be reduced to an absolute minimum and done in small mobile vans or at fixed sights?
7. **Source data systems need integration.** The timeliness and quality of reporting of drug overdose deaths would be improved drastically if we developed the capability of the systems used by our data providers to share data instead of developing another disease-specific reporting system. In this case, the ability of data to be passed automatically between medical examiner/coroner systems, electronic death registration systems and electronic health records would not only improve the timeliness, robustness, and quality of drug overdose deaths reporting, but also make us ready to report on the next public health crisis.
8. **This challenge is unique to NCHS. The National Death Index (NDI) needs to be funded through appropriations and not self–supported through charges to the researcher.** Since the funds collected for the NDI go back to the states for their vital

THE AMERICAN STATISTICIAN HIGHLIGHTS

August Issue Spans Methodology, Applications

Daniel Jeske, *The American Statistician* Executive Editor

The August 2018 issue of *The American Statistician* features eight articles and two letters to the editor that span a wide range of interesting methodology and application areas. As usual, there is something for everyone in this issue.

The General section begins with an article about the construction of minimum volume confidence sets for the parameters of a shifted exponential distribution (also known as the two-parameter exponential distribution). A second article proposes a new test for detecting general forms of serial temporal dependence.

We have three articles in the Statistical Practice section. First, an article presents a structural equation modeling approach for analyzing a spatial regression model for situations in which there is spatial influence on both the response and covariate variables. The second article in this section presents a semiparametric Bayesian model for analyzing homerun production of Major League Baseball players. The final article details a random graph model for studying citation patterns within the causal inference literature.

We have one article in the Teacher's Corner that focuses on a variety of thought-provoking insights on the Wilcoxon-Mann-Whitney test. In the Short Technical Note section, you will find a probabilistic proof of an interesting binomial coefficient identity. The Interdisciplinary section is represented by an in-depth analysis of if and how the proportion of gun-related suicides can estimate gun prevalence.

The issue concludes with two letters to the editor that provide comments about a recent article that investigated the use of interpolated nonparametric confidence intervals for population quantiles.

To read these and other articles in *The American Statistician*, visit *https://bit.ly/2uIbVIf.*

statistics systems, appropriated dollars should go to the states to support these systems so needed by the federal government. In a similar vein, we need to consider a National Birth Index (NBI) to be able to better assess infant and child health programs. With an NBI, electronic health records, and the NDI, we would have a framework for a health registry system for the US.

9. **Data linkage.** We need to drastically improve and expand our data linkage activities with administrative data sets within HHS (US Department of Health and Human Services) and across departments.

10. **Cybersecurity.** Cybersecurity needs to be improved for all agencies, including NCHS, and it will be very expensive to initiate. Initially, I believe it will unfortunately become the main stumbling block to innovation in making our data more available. In the short-run, cybersecurity will cause us in government to become even more risk adverse than we are now … hard to believe! If we are going to make data more available—if we are going to link administrative and survey data across agencies as we did when we linked HUD (US Department of Housing and Urban Development) data with NHANES data and showed a positive impact on blood lead levels of children whose families are receiving HUD support—then we need a safe place to do that. At this point, that safe place needs—dare I say it—to not be connected to the networked world we now work in! Thus, research data centers (RDCs) and the FSRDC (federal statistical research data centers) at universities may be a good first and safe step while we develop other more flexible methods. This issue was brought up in the recent report of the Commission on Evidence-Based Policymaking, which argued for a National Secure Data Service.

Am I dismayed? Is this just an old man's lament about wanting things to go back to how they were? Absolutely not. I wish I were 40 years younger with the technical tools and knowledge to take this on. We can do this, we really can! We in federal statistical agencies want to change how we collect and disseminate our data. You either need to invest in us so we can change or keep us going as we are and invest in new organizations that will do what we are doing in different modalities. Whatever is decided, an investment is needed.

I am encouraged because the president's management agenda has a major goal of leveraging data as a strategic asset. As I understand it, this cross-departmental effort will be to provide examples of secure data sharing, developing standards for combining data, implementing the recommendations coming from the Commission on Evidence-Based Policymaking, strengthening the FSRDCs, and developing common terminology—which will be so necessary to future data-linkage activities.

Just last week I was at a conference sponsored by OMB (Office of Management and Budget) on leveraging data as a strategic asset and I was impressed with the energy and interest in this administration to improve our ability to strategically collect, use, and share data across government and with the research community. This is a great time to shape the future of federal statistical agencies. ■

Nontraditional Collaboration: A Success Story at JSM 2016

We have all heard the saying, "Teamwork makes the dream work." Teamwork, or collaboration, is a key component of success in any business, and academic research is no exception.

Traditional academic collaborations are often intentional, formed by a defined team with a shared purpose and well-established rules and strategies. On the other end of the spectrum is the nontraditional collaboration. These are accidental and often emerge out of casual conversations in random places and, occasionally, out of curiosity. No matter what type it is, when a collaboration is successful, it produces a fruit all parties find edible.

One such fruitful (nontraditional) collaboration came together at JSM 2016 when a group of five early-stage statisticians decided to meet up over a cup of coffee with the sole purpose of reconnecting after a gap of more than six years. The team was comprised of three recent graduates (**Broti Garai**, **Prithish Banerjee**, and **Shrabanti Chowdhury**), one graduate student (**Saptarshi Chatterjee**), and one first-year postdoc (**Himel Mallick**).

This is an unusual collaboration strategy for young researchers to undertake, and they faced several barriers simultaneously. First and foremost, publishing in a top-tier statistical journal without an expert senior author is difficult. Second, there was no major funding support. Third, all project members had full-time commitments, mandating project duties only during nights and weekends.

Despite all fears, the team decided to take the route less traveled and worked for more than a year to achieve the short-term and long-term goals of the project. "Although the progress was slow, it was not surprising given the nature of the collaboration, said Mallick. "Not having a formal deadline from a funding agency helped, so did the free availability of primary analytical and collaborative tools such as R, Slack, Dropbox, Google Docs, and GitHub, among others."

The project faced additional obstacles when it came to publishing the research findings. A recent *Proceedings of the National Academy of Sciences* study (see *https://bit.ly/2uh0act7*) concluded that single-blind reviewers are significantly more likely than their double-blind counterparts to recommend for acceptance from famous authors, top universities, and top companies. The experience was somewhat similar for the accompanying manuscript without a renowned name in the author list.

Lessons Learned

As young researchers, it is certainly possible to successfully collaborate without an expert onboard. One should not underestimate the value of a casual conversation, especially in a venue like JSM, which can often (accidentally) lead to a successful collaboration. As an early-career scientist and statistician, it might be difficult to publish in top-tier statistical journals without a senior author, but it is not impossible. It is important to be open-minded about nontraditional collaboration opportunities, which can have rocky beginnings but happy endings.

Finally, there is one aspect that is common between nontraditional and traditional collaborations. A successful collaboration is not defined by the cool science only. Clearly stipulated roles for individual team members and a specified approach toward achieving the goal are as important as the science itself. It is, therefore, important to choose collaborators wisely and think carefully before making a commitment.

The initial draft was rejected from all the journals the team submitted it to. After surviving a few rounds of peer-review, the paper finally saw the light of the day when it was accepted by *Statistics in Medicine*, almost two months ago now. The published version of the paper is available at *https://bit.ly/2JhjYRA*.

But that didn't happen without another setback the team had no experience dealing with. *Statistics in Medicine* initially rejected the paper despite positive comments from the reviewers. The team decided not to get bogged down by the rejection, but instead challenged the editor's decision. It was a particularly daunting task for a young team to consider. "We did not think we would be successful, but we took a shot at the rebuttal letter," said Mallick. "And to our surprise, our resubmission request was approved and the journey was smooth from that point onwards."

The team took an unconventional approach to determining authorship credit. Co-first authorship is a tricky business in general and certainly not common in biostatistics. "We decided to break with tradition and have co-first authorship for all three authors who put in an equal amount of intellectual contribution," said Mallick. "That was the most fair thing to do." ■


Garai


Banerjee


Chowdhury


Chatterjee


Mallick

Why Would USDA Shoot Itself in the Foot?

Katherine Smith Evans

Katherine Smith Evans was the administrator of USDA's Economic Research Service from 2007–2011.

The US Department of Agriculture (USDA) proposed to substantially cut funding for its policy research agency, the Economic Research Service (ERS). ERS products inform far-reaching program and policy decisions in the farm, food, and related sectors—sectors that together comprise 5.5 percent of US gross domestic product (GDP), more than half the country's land and resource base, 46 million rural Americans, and more than 40 million low-income Americans unable to afford a nutritionally adequate diet. ERS has an outsized impact by directly serving USDA programs and a broad swath of Americans.

While the House and Senate appropriations committees propose to keep ERS at its 2018 funding level in 2019, it is still useful to note that the administration proposed to throw ERS under the bus, cutting its funding by 48 percent in 2019.

You may know ERS for its food price increase projections, farm income figures, and agricultural productivity measures. But the part of its program President Trump's administration would cut is its objective, data-driven, timely, and robust analysis—analysis that makes USDA programs more cost-effective.

ERS has served Republican and Democratic administrations equally well by documenting the economic consequences of different options for the farm, food assistance, agricultural trade, natural resource, and rural development programs the USDA runs. In so doing, it presents efficient program management alternatives.

For example, ERS developed an indexing system that helped USDA get more soil conservation for each dollar spent. This work wouldn't be supported under the administration's proposal.

Trade negotiators have relied heavily on ERS' quick-turnaround analysis of particular agricultural trade scenarios to decide the best US stance on trade agreement possibilities. Again, this would be cut under the administration's proposal.

ERS' interactive, continuously updated Atlas of Rural and Small Town America, which measures rural areas along more than 85 dimensions, is a tool heavily used by community and business leaders to help communities thrive in the global economy. But the Trump budget directs ERS to discontinue all research and statistics related to the rural economy.

ERS' work on the effectiveness of USDA nutrition assistance programs was highlighted as sound, fact-based evaluation by the congressionally established bipartisan Commission on Evidence-Based Policymaking. ERS conducts research on access and barriers to affordable, healthy food and the consequences of limited food access on food spending, diet, and health—activities that would not be supported by the administration's 2019 proposal. ERS conducts studies and evaluations of the Supplemental Nutrition Assistance Program (SNAP), focusing on SNAP's effectiveness in supporting income and diet quality objectives, and evaluates the Special Supplemental Nutrition Program for Women, Infants, and Children; National School Lunch; and other food assistance programs that have guided program decisions in USDA's Food and Nutrition Service.

But the Trump budget would explicitly "eliminate research and other data products on USDA food and nutrition assistance programs and on food consumption and nutrition, including all data resources related to food access and consumer food choices, including … Food Access Research … the SNAP Data System, Fruit and Vegetable Prices, and Price Spreads from Farm to Consumer." The FY 2019 administration's budget would not even support the extensive and unique Food Acquisitions and Purchases Survey conducted by ERS, which is the statistical basis for numerous studies inside and outside of USDA. No other public or private entity has or collects data that precisely connects food prices, type, and quantity of food purchases (both for consumption at and away from home); household purchasing power; and food provided via food assistance programs and food banks.

In a move that suggests a lack of knowledge about how ERS research and economic indicators rely on statistics, the administration would direct ERS to continue its work to track and project farm income, but would specifically cut $1.65 million from the Agricultural Resource Management Survey (ARMS). The ARMS provides the basic financial accounting information from farm businesses and farm households that critically underpins sound farm income and wealth analysis.

Fortunately, Congress is not going along with this move. Agricultural appropriations leaders Sen. John Hoeven, Sen. Jeff Merkley, Rep. Robert Aderholt, and Rep. Sanford Bishop understand that ERS, with its statistical collections and subject-matter expertise, is not reproducible elsewhere.

Still, it remains a puzzle why the USDA wouldn't want to be able to evaluate its own nutrition assistance, trade, rural development, and other critical programs. Why would it disregard the implications of trade policy for farmers, food businesses, and consumers; the state of rural Americans and institutions; and documentation of how and how well USDA programs achieve their goals? The failure to fully support the ERS is short-sighted and indicates a disregard for evidence-based policymaking in the USDA. ■

STAT*tr@k*

Lessons on Leadership, Influence, and How to See the Big Picture from My Career

I was invited to give a talk in the JSM 2016 short course, "Preparing Statisticians for Leadership: How to See the Big Picture and Have More Influence." My first reaction was, "How can I teach about leadership and influence? I've never had any formal training!" But the organizers of the course assured me that if I could talk about lessons on leadership and influence I had learned from my career experiences, it would be informative to the students. So, I thought about a few situations in which I had been in the position to lead and influence others and, from those "case studies," I inferred some principles and practices. This STAT*tr@k* article is based on the JSM 2016 talk and an "encore" presentation I was invited to give at JSM 2017.

The case studies to be presented cover three types of activities from my career: collaborating on a statistical application at the National Center for Health Statistics (NCHS); developing the Conference on Statistical Practice as a volunteer for the ASA; and a foray into management as director of the Division of Research and Methodology at NCHS. By the way, the diversity of these activities illustrates you don't need to be a manager or boss to lead and influence. Indeed, I shied away from management for most of my career, serving as a mathematical statistician at the US Census Bureau from 1985–1988, a faculty member in biostatistics at UCLA from 1988–1999, and a senior research scientist at NCHS from 1999–2010. It wasn't until "my arm was twisted" that I began as director of the Division of Research and Methodology at NCHS in 2010.

Of course, getting into management provides one type of opportunity to lead and influence. And managers often "have the ears" of higher ups in the organization, which helps to get things done. But statisticians have many other opportunities to lead and influence, such as getting involved in important projects and steering them in the right direction, teaching, and conducting influential research. So, get involved in things you care about and do your best. Chances are, you'll lead and influence.

Multiple Imputation of Missing Income Data

One of my biggest projects at NCHS involved missing income data in the National Health Interview Survey (NHIS), which is the principal source of information about the health of the noninstitutionalized US population (*https://bit.ly/2NN8NDt*). The NHIS also collects a variety of socioeconomic and other data, which analysts often relate to the health information.

Like many surveys, the NHIS has had relatively high levels of missing data on income and earnings. For example, in the late 1990s and early 2000s, data on family income were completely missing or only partially complete about 30% of the time. Beginning with the 1997 survey year, NCHS has multiply imputed missing income data in the NHIS annually. The development and evaluation of the

Nathaniel Schenker earned a PhD in statistics from The University of Chicago. He worked at the US Census Bureau, University of California at Los Angeles, and National Center for Health Statistics before retiring last year. He has been very involved in service to the profession and was the ASA president in 2014.

Lessons Learned

Each of these lessons applies to a variety of leadership activities.

1. Communication is crucial.
2. Find out and address what is needed.
3. Emphasize benefits and value.
4. Lead with energy and enthusiasm.
5. Be clear about downsides as well as upsides.
6. If applicable, obtain support from higher-ups.
7. Ensure continuing success.
8. Let others shine.

imputation methods are discussed in the 2006 *Journal of the American Statistical Association* (*JASA*) article, "Multiple Imputation of Missing Income Data in the National Health Interview Survey."

When I worked on this project, I was a senior research scientist in the Division of Research and Methodology, which is NCHS's methodological research, development, and consulting division. I was collaborating with staff from other divisions, including the Division of Health Interview Statistics (DHIS), which was responsible for running the NHIS.

Challenges

Besides the relatively high nonresponse rates on income data, there were other complexities that made the project challenging but also statistically interesting. For example, the data structure was hierarchical, with income measured at the family level but earnings measured at the person level. Moreover, family income could be reported as an exact value or as falling in a range of values. And personal earnings could only exist if a person was employed, but sometimes employment status was missing, as well.

Along with the statistical challenges, there were organizational challenges. For example, DHIS had a tight production schedule and processes already in place that worked well. Introducing new statistical methods into the processes could be risky. Moreover, neither DHIS staff nor the ultimate internal and external analysts of the data were overly familiar with multiple imputation.

Keys to Success

Before starting the multiple-imputation project, I needed to communicate to my colleagues from DHIS both the potential value multiple imputation of income data could provide and possible downsides of the approach. I also needed to communicate the pros and cons of the various alternatives available. Such communication helped ensure the DHIS staff (my "clients") were involved in the decision process and felt well-informed about what was being done. Importantly, the director of DHIS supported the project, which helped it move forward.

Another key to success was ensuring the multiple-imputation process could be carried out within DHIS. We needed to develop a system, in collaboration with DHIS staff, that could be inserted into the division's existing processes and used software with which the division staff were familiar.

Because SAS was the predominant software at NCHS, we developed a system that used SAS and SAS-callable IVEware, and we developed approximations when needed to make the imputation feasible. Finally, to assist analysts who would be using the multiply imputed data and who might not be familiar with the technique, we developed sample analysis software.

Rewards

The multiple-imputation system helped promote good science both within the agency and among outside data analysts. Because we developed a system that was relatively understandable, in close collaboration with DHIS, it could be applied by the division to subsequent years of the survey with minimal consultation. Finally, the project team received an NCHS Director's Award for their work and coauthorship on an article in *JASA*.

Developing the ASA's Conference on Statistical Practice

In 2008, when I was an ASA vice president, the president-elect appointed me to chair a workgroup on meetings charged with considering ways to expand the ASA's meetings portfolio to increase revenues and provide value to members. Based on ideas developed by this workgroup and through the work of subsequent committees, the Conference on Statistical Practice (CSP) was established (see *https://bit.ly/1zjQtUK*).

Challenges

Was a new conference needed? If so, what type? Once those questions were answered, we needed to sell the idea to the membership, ASA staff, and ASA Board of Directors. And we needed to avoid stepping on the toes of other ASA groups that already had conferences. Finally, we needed to make sure the conference would get off the ground.

Keys to Success

The workgroup reviewed the existing statistical conferences, both within and outside of the ASA, to see what gaps there might be. We studied the history of

the ASA's former winter meetings to determine their positive and negative aspects and why they ended. We also listened to the membership. In particular, it seemed clear that applied statisticians wanted more from the ASA.

Once we decided to propose a conference for applied statisticians, we presented our ideas to a number of ASA groups, obtained and responded to their feedback, and worked to involve representatives of those groups in planning the conference. I also gave regular progress reports to the ASA Board and followed up on their suggestions. These efforts helped ensure the planning and establishment of the proposed conference proceeded without too much objection.

Of course, any idea—however good—needs follow-through to succeed. In the case of CSP, the ASA appointed committees to flesh out and enhance the workgroup's ideas and create and implement plans for the conference.

Rewards

The CSP has been well-received. Reviews of it have been positive, and attendance has grown since the first conference in 2012. Although it is not a big money-maker yet, the CSP provides value to an important component of our profession—practicing statisticians. Providing such value is crucial to both the ASA and the profession. My workgroup and the succeeding CSP committees have taken pride in being able to promote the CSP's goals and success.

My Foray into Management

In 2010, at the request of the director of NCHS, I finally tried my hand at management by becoming the director of the Division of Research and Methodology (DRM), which—as mentioned earlier—is NCHS's methodological research, development, and consulting division. DRM comprised an office of the director and three branches: Collaborating Center for Statistical Research and Survey Design; Collaborating Center for Questionnaire Design and Evaluation Research; and NCHS Research Data Center. The total size of DRM was about 50 staff members.

Challenges

When I began as director, I thought staff morale could use some enhancement. There had been a substantial amount of staff attrition, so additional hiring was needed. I wanted to change the research culture somewhat by encouraging the staff to publish their work in peer-reviewed outlets. Finally, I wanted to increase the integration of DRM with the rest of NCHS.

Keys to Success

Before taking on this management position, I wrote down my vision, ideas, and requests for the division and obtained agreement from the director of NCHS to ensure I'd have the best chance of accomplishing my goals.

Communication with and visibility to the DRM staff were important for morale. I held quarterly staff meetings, updated the staff on important happenings, and made myself as available as possible to meet individually with staff members when needed. I expressed my vision for DRM to be part of a team (NCHS) doing important work, with DRM's staff conducting research relevant to and motivated by the work of the agency. With regard to increasing the size of the staff, I put a lot of thought and effort into hiring and did my best to support my branch chiefs in their hiring efforts, with an emphasis on hiring high-quality staff members.

To encourage the staff to publish their work in peer-reviewed outlets, I stressed the role of peer review as a foundation of good science. I emphasized the value of peer-reviewed publications in enhancing the status of NCHS and DRM. Moreover, I let the staff know that, although I didn't want to lose any of them, publishing their work in peer-reviewed outlets would make them more marketable. Finally, I did my best to provide staff with time to work on basic research, with the provision that it should be relevant to the agency's work—at least in the long term.

To increase DRM's integration with the rest of NCHS, I tried to "make friends" throughout the agency, for example, by inviting other division directors to lunch, asking about their needs, and encouraging them to collaborate with DRM on projects. DRM also publicized its work relevant to the agency.

Rewards

Frankly, I didn't find management to be rewarding in and of itself. However, I did find the results of my efforts rewarding. The morale within DRM seemed to increase, and we were able to hire several new, talented staff members. The number of peer-reviewed publications increased. We contributed to NCHS's mission to provide statistical information that guides actions and policies to improve the health of the American people. We received appreciation from elsewhere in the agency and also had a positive external review. Finally, I enjoyed and celebrated the accomplishments of my staff and espoused the philosophy, "Let your staff members shine, and it will reflect well on you."

A Shout-Out to Teaching

If I had the space, I would include a fourth case study, namely teaching introductory biostatistics at UCLA. Teachers lead students in learning, and such leadership can be especially challenging in required courses for non-majors. Many of the keys to success for leadership and influence in other endeavors—such as those presented here—are important for teaching, as well. ■

PASTIMES OF STATISTICIANS

What Does Andrew Althouse Like to Do When He Is Not Being a Statistician?

Andrew Althouse

Who are you, and what is your statistics position?

My name is Andrew Althouse. I am a biostatistician at the University of Pittsburgh, with a faculty appointment in the department of medicine, supporting clinical research initiatives in a variety of medical specialties. I have collaborated with physicians and written papers about diabetes care, cardiology, cardiac surgery, obstetrics, and gynecology—just to name a few!

Tell us about what you like to do for fun when you are not being a statistician.

I enjoy watching and participating in strength and conditioning sports: power lifting, weightlifting, CrossFit, and Strongman competitions. I perform some type of "strength workout" four to six days per week. My training incorporates several tools—mostly heavy barbell lifts (back squats, front squats, and deadlifts), but I also use kettlebells (a big iron ball with a handle on it), heavy sandbag carries and lifts, and other odd-object lifting whenever possible. I exercise for about 30 minutes every night when I get home from the office … and usually unwind with a nice cold beer afterward!

What drew you to this hobby, and what keeps you interested?

I started lifting weights as a teenager with my father, who competed as a Masters powerlifter for several years (and still lifts weights). I carried this hobby/passion on through high school and college as a football player, and when I purchased my first home last year, I couldn't wait to outfit a "garage gym" setup with a squat rack, barbells, and bumper plates from Rogue Fitness.

Althouse lifts 585 pounds. His heaviest lift is 600 pounds, but 585 pounds is the biggest lift he has completed in his garage.

What keeps me interested? There's both a physical component and a mental component to my enjoyment of lifting weights. Physically, this hobby keeps me feeling healthy, strong, and vibrant. Mentally, the challenge of continuing to make progress and the self-satisfaction one gets from completing good lifts are gratifying. There is perhaps no other hobby where one can draw such a direct line from "work done" to "progress made." Academia and professional life often feel like a difficult series of ups and downs. ■

STATS4GOOD

STAT*tr@k* and Data for Good: A Perfect Combo

Membership in the American Statistical Association offers many benefits: chapter membership, education and collaboration opportunities, great magazines, and much more. One of the most important resources for students and people just getting started in a career in statistics is STAT*tr@k*. Describing itself as "a website produced by the American Statistical Association for individuals who are in a statistics program, recently graduated from a statistics program, or who recently entered the job world," it is much more than a website. Under the care and guidance of ASA staff, STAT*tr@k* provides access to an extensive collection of resources for the early-career statistician.

Content and Resources

From mentoring programs to articles with valuable career advice, to information about applying for scholarships and fellowships, to educational opportunities and students in the news, STAT*tr@k* is a hub for early-career development information. You can find out about hackathons and learn from others' experiences to improve your performance; get the skinny on conferences and what they offer for early-career development; and learn about the work of local chapters, ASA sections, and other organizations such as the National Institute of Statistical Sciences (NISS). With so much information and resources available, STAT*tr@k*'s search function might be the most important part.

One of the most valuable aspects of STAT*tr@k* is the opportunity to gain writing experience and exposure by submitting articles. You can share your story, experiences, opportunities, and useful information for students and early-career statisticians. For example, I had the opportunity to interview Megan Price from the Human Rights Data Analysis Group for the September issue of *Amstat News*, which focuses on careers in statistics. Because our talk focused on her career as a human rights statistician, this interview will appear on STAT*tr@k*, as well.

STAT*tr@k*'s support for people early in their career makes it a resource for the entire statistical community, so everyone can participate. It's a perfect channel for senior statisticians and leaders in the statistical community to share experiences, resources, and opportunities with others just getting started.

STAT*tr@k* and Data for Good

Looking to get started in Data for Good projects? STATtr@k is a great place to research opportunities, learn about what others are doing, and connect with the project that meets your interests and develops your skills. It's also a great place to let people know about your Data for Good projects and recruit newcomers to your important work. Using the search function (it's in the upper right corner of every page), look for the subjects and opportunities that interest you most.

For example, a search on "social good" will turn up information about scholarships, student fellowships, new undergraduate programs in data science, and an article about how internships helped four students make a difference. A search on "justice" will connect you with volunteer opportunities, the ASA's Research Experience for Undergraduates (REU) program, best practices for developing an open data portal, and feature articles like "The Local ASA Chapter Is My Justice League" (*https://bit.ly/2JlGoRB*), in which Scott McClintock described his ASA chapter—in Philadelphia—as his "Justice League," where collaborators use their "statistical super-heroism for the greater good." The list of resources and possibilities for good work goes on and on. If you don't find information about a program, it means you can write about it and let STAT*tr@k* staff know so others can benefit from your experience.

The mission of the ASA can be summed up as doing good statistics, doing good for statistics, and doing good *with* statistics. The wealth of resources STAT*tr@k* offers early-career statisticians, used to support Data for Good projects, meets these objectives at once. My highest hope for Data for Good, in this column and elsewhere, is to see it become normative—a natural, ordinary part of a career in statistics. That starts from the beginning, and so naturally connects with STAT*tr@k*'s mission to provide resources for early-career statisticians.

We all have a role to play in developing the statistics—and the statisticians—of the future. That means we all can be users, even contributors, to STAT*tr@k*'s work to support students, recent graduates, and others just starting out. Help make Data for Good an important part of their statistical career. Visit *http://stattrak.amstat.org* and find your place in moving statistical science forward as a powerful means for doing good in our society, communities, and world. ■

STAT*tr@k* 411

Website: *http://stattrak.amstat.org*

Includes sections for awards and scholarships, getting in touch with ASA chapters, career support, and resources

Updated monthly

Search function, including archives

Submit articles here: *https://bit.ly/2NcALaM*

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

ASA photo/Meg Ruyle
Barry Nussbaum, 2017 ASA president, gives the banquet talk, "I Never Met a Datum I Didn't Like."

Symposium on Data Science and Statistics: A Remarkable Success

Yasmin H. Said, 2018 SDSS Program Chair

Yasmin H. Said

More than 500 people attended the sold-out 2018 Symposium on Data Science and Statistics in Reston, Virginia, May 16–19.

The program for this first ASA symposium featured a strong program offering short courses, concurrent sessions, and electronic poster sessions. There also was an exhibit hall and many opportunities for networking. Emery N. Brown gave the keynote address "Uncovering the Mechanisms of General Anesthesia: Where Neuroscience Meets Statistics," while David Scott, Adalbert Wilhelm, and Jerome Friedman each gave a plenary talk.

Photo courtesy of Laura Normoyle Photography
From left: 2017 ASA President Barry D. Nussbaum; Ross McKitrick from University of Guelph in Canada; Emery N. Brown from MIT, Harvard Medical School, and Massachusetts General Hospital; and SDSS 2018 Program Chair Yasmin H. Said from George Mason University

Within the invited program were sessions on data science, data visualization, machine learning, computational statistics, computing science, and applications—some standing room only. The short courses, which took place May 16, also were full.

One of the most popular invited sessions was "Interactive Statistical Graphics: Where Are We Now?" It featured talks by Wayne Oldford ("Exploratory Visualization via Extendible Interactive Graphics"), Catherine Hurley ("Model Exploration via Conditional Visualization"), and Heike Hofmann ("Interactive Web-Graphics Using R").

Photo courtesy of Laura Normoyle Photography

From left: Katherine Ensor from Rice University, Alyson Wilson from North Carolina State University, Dave Higdon from Virginia Tech, and 2006 ASA President Sallie Keller from Virginia Tech

ASA photo/Meg Ruyle

Martha Christino from TC Williams High School presents her poster on Hurricane Irma to Peter Craigmile, a professor at The Ohio State University.

Some other talks garnering packed rooms included the following:

- Daniele Struppa: "Social Networks and Simplicial Complexes"
- Menas C. Kafatos: "Laws of the Universe, Information, and Mind in the Quantum Universe"
- Kirk Borne: "Exploring and Exploiting Interestingness in Data Science"
- Leland Wilkinson: "Automatic Visualization"
- David Banks: "Cherry-Picking Techniques for Complex Data Sets"
- Edward George: "Bayesian Penalty Mixing with the Spike and Slab Lasso"

To read about all the sessions and talks, visit the online program at *https://bit.ly/2NMG4yJ*.

The banquet talk, "I Never Met a Datum I Didn't Like," was given by Barry D. Nussbaum, the 112th president of the American Statistical Association and chief statistician for the US Environmental Protection Agency. At the banquet, the Interface Foundation of North America and ASA awarded a lifetime achievement award to Edward J. Wegman for his seminal contributions to computational statistics, data visualization, and data science. In 1987, he incorporated the Interface Foundation and has been the treasurer for 31 years.

Photo courtesy of Laura Normoyle Photography

Edward George from the University of Pennsylvania, center; Hamparsum Bozdogan of the University of Tennessee, center right; and Aylin Alin from Dokuz Eylul University in Turkey, right, attend a session at SDSS 2018. Behind them and to the left is Michael Schimek from Medical University of Graz, Austria.

Keynote and Plenary Speakers

Emery N. Brown is a renowned scholar and member of the National Academy of Medicine, National Academy of Sciences, and National Academy of Engineering. He is an anesthesiologist-statistician whose experimental research has made important contributions to understanding how anesthetics act in the brain. In his statistics research, he has developed signal processing algorithms to study dynamic processes in neuroscience.

David Scott is the Noah Harding Professor of Statistics at Rice University in Houston, Texas. He was a founding member of the department of statistics in 1987 and its chair. Scott's talk focused on Edward Wegman's influence on the profession and his work in computational statistics and density estimation.

Adalbert Wilhelm holds a professorship in statistics and is the vice dean of the Bremen International Graduate School of Social Sciences at Jacobs University in Bremen, Germany. His talk focused on statistical graphics in data science. He bridged the different visualization aspects from computer science, statistics, and application domains and discussed recent trends.

Jerome Friedman is a renowned scholar and member of the National Academy of Sciences. He is a professor of statistics at Stanford University and one of the world's leading researchers in statistics and data mining. Friedman's talk was titled, "Omnibus Regression: Predicting Probability Distributions with Imperfect Data."

This symposium is a continuation of the Interface Symposium on Computing Science and Statistics. The first Interface Symposium was held in Reston, Virginia, in 1988. ■

2018 POSTER & PROJECT COMPETITIONS

The American Statistical Association is pleased to announce the winners of the 2018 Poster Competition (newly renamed the ASA Data Visualization Poster Competition) and Project Competition.

First-place winners received $300, a plaque, a plaque for their school, and grade-appropriate graphing calculators for the students and advisers provided by Texas Instruments. Second-place winners received $200 and a plaque; third-place winners received $100 and a plaque; and honorable mentions received plaques.

The poster and project competitions are directed by the ASA/NCTM Joint Committee on Curriculum in Statistics and Probability. The 2018 ASA Data Visualization Poster Competition leader was Rodney Jee of Discover Financial Services. Nathan Kidwell of Oaxaca Christian School in Oaxaca, Mexico, served as the head project competition leader.

As well as a name change, 2018 also brought a new rubric for the poster competition (see *https://bit.ly/2LcpI0F*). K–12 posters are due every year on April 1. Projects (written reports) for grades 7–12 are due every year on June 1. Visit the competitions web page at *https://bit.ly/2a1DpkO* for information—including previous winners, entry forms, instructional webinars, and the rubrics used for judging the posters and projects.

2018 Regional Poster Competition Leaders

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

Connecticut Chapter Statistical Poster Competition

Valerie Nazzaro, Wesleyan University
Jennifer McGinniss, Sunovion
https://bit.ly/2us9Olc

Kansas/Western Missouri Statistics Poster Contest

Ananda Jayawardhana, Pittsburg State University
https://bit.ly/2ulg2uq

Michigan Statistics Poster Competition

Dan Adrian, Grand Valley State University
https://bit.ly/2Nf6ZC6

Nevada K–12 Statistics Poster Competition

Tia Price, Southern Nevada Regional Professional Development Program
https://bit.ly/2NNGa9e

Ohio Statistics Poster Competition

Jerry Moreno, John Carroll University
https://bit.ly/2NO7Fj5

Pennsylvania Statistics Poster Competition

Pete Skoner, Saint Francis University
https://bit.ly/2Lh0RJb

Pullman, Washington Statistics Poster Competition

Dean Johnson, Washington State University
dean_johnson@wsu.edu

Washington Statistical Society Poster Competition (DC Metro Area)

Elizabeth Petraglia, Westat
https://bit.ly/2LbGIHw

ASA National Poster Competition

Leader: Rodney Jee, Discover Financial Services

Contact: Rebecca Nichols, ASA Director of Education, *rebecca@amstat.org*
https://bit.ly/2LbGIHw

GET INVOLVED

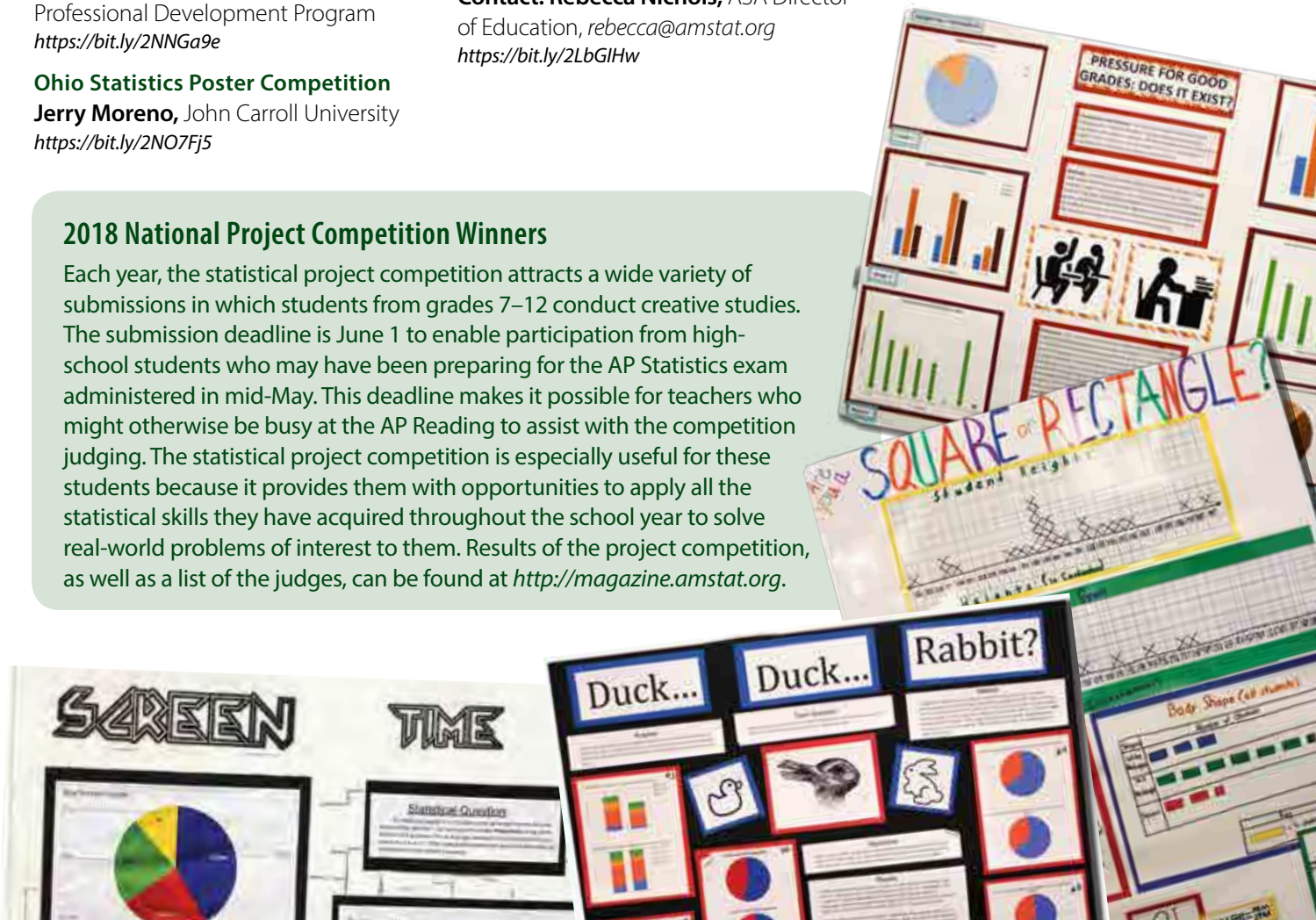
For information about how you can start a regional poster competition or mentor students in your area, see the article appearing in the July 2011 issue of *Amstat News* at *https://bit.ly/2NN6TCZ*.

You can download a flier about the ASA poster and project competitions and other K–12 statistics education programs and resources to share with your local schools at *https://bit.ly/2zCC9S4*.

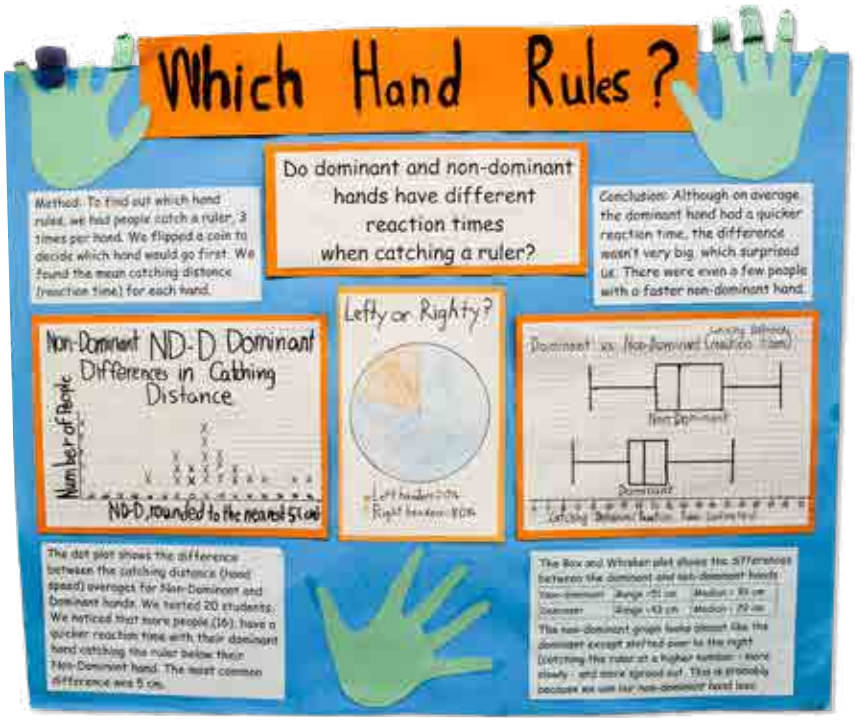
For additional information or questions regarding how to get involved in the poster or project competitions, contact ASA Director of Education Rebecca Nichols at *rebecca@amstat.org*.

2018 National Project Competition Winners

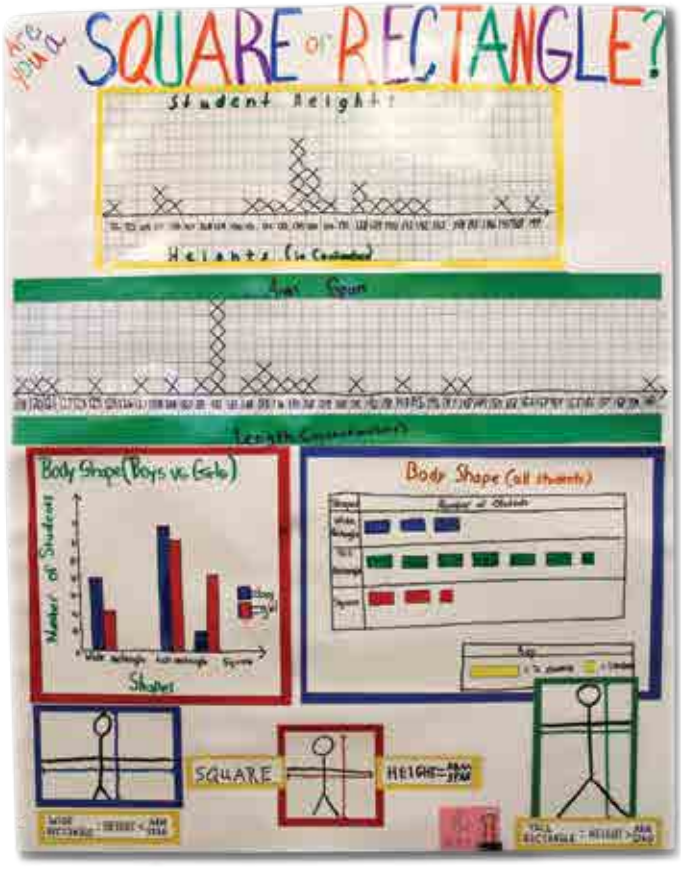
Each year, the statistical project competition attracts a wide variety of submissions in which students from grades 7–12 conduct creative studies. The submission deadline is June 1 to enable participation from high-school students who may have been preparing for the AP Statistics exam administered in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



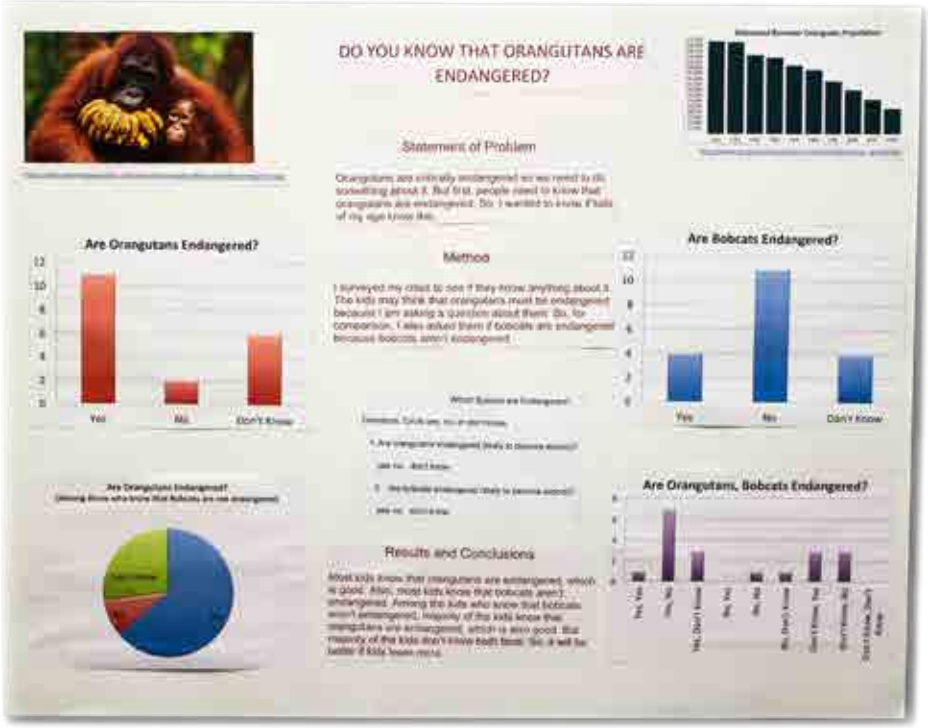
POSTER COMPETITION WINNERS GRADES K–3



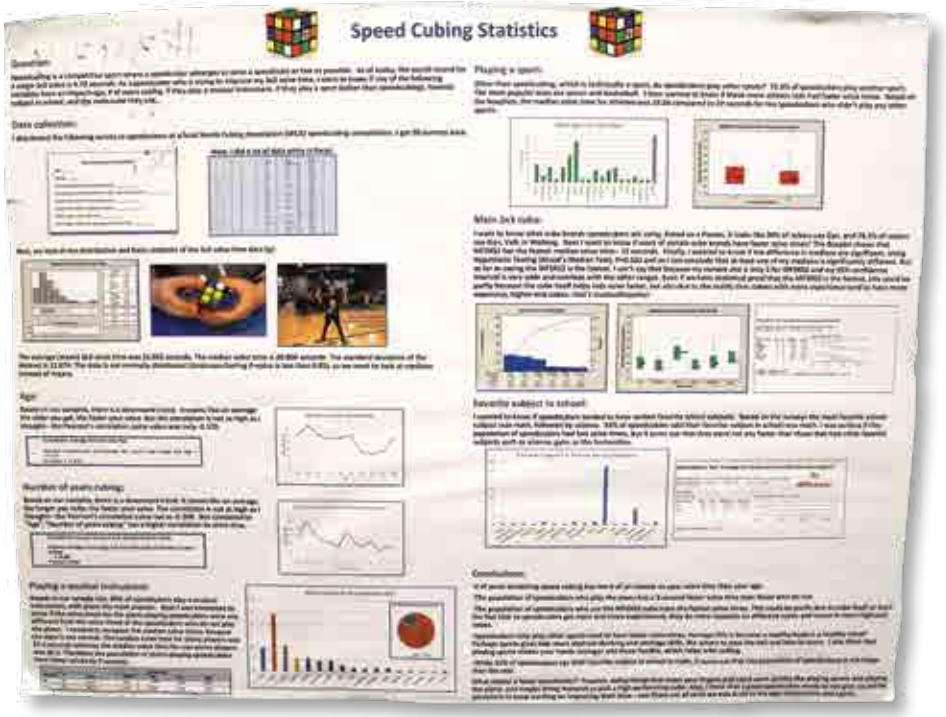
Alex Ren, Julia Ding, and Cora VanDongen



Third-Grade Class at Timmons Elementary School in Chagrin Falls, Ohio



Arushi Choudhary



Salar Nasimov

FIRST PLACE

Alex Ren, Julia Ding, and Cora VanDongen
"Which Hand Rules?"
Conners Emerson School
Bar Harbor, Maine

SECOND PLACE

Third-Grade Class
"Are You a Square or a Rectangle?"
Timmons Elementary School
Chagrin Falls, Ohio

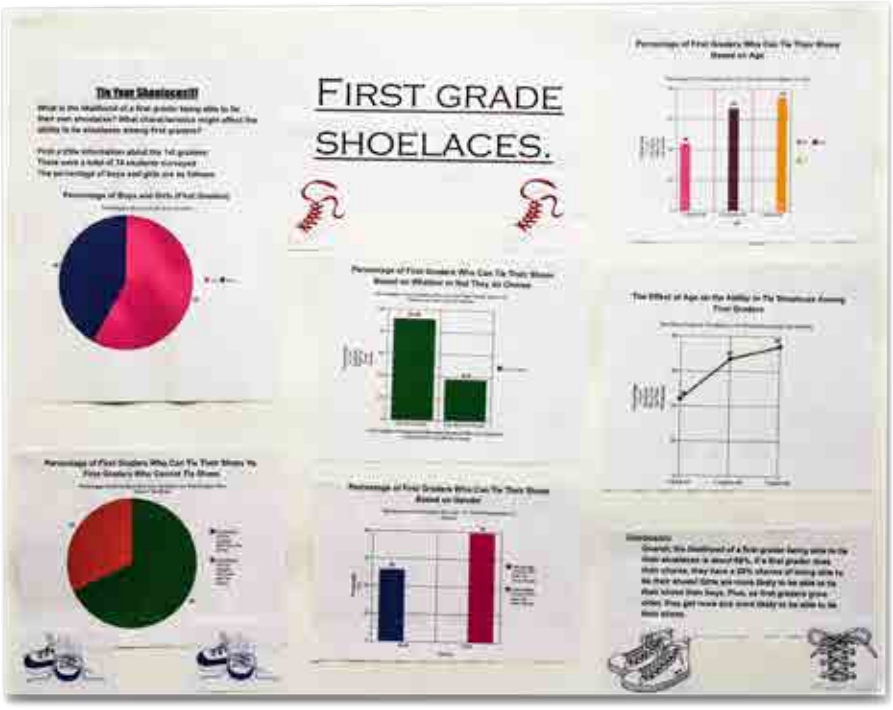
THIRD PLACE

Arushi Choudhary
"Do You Know That Orangutans Are Endangered?"
Skaggs Elementary
Plano, Texas

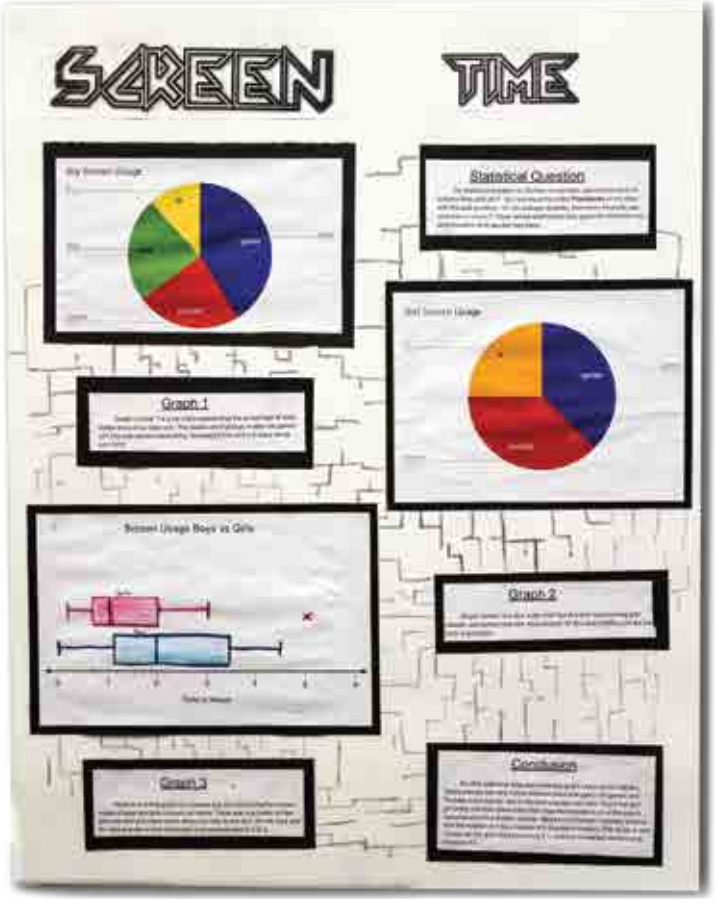
Honorable Mention

Salar Nasimov
"Speed Cubing Statistics"
Fairhill Elementary
Fairfax, Virginia

POSTER COMPETITION WINNERS GRADES 4–6

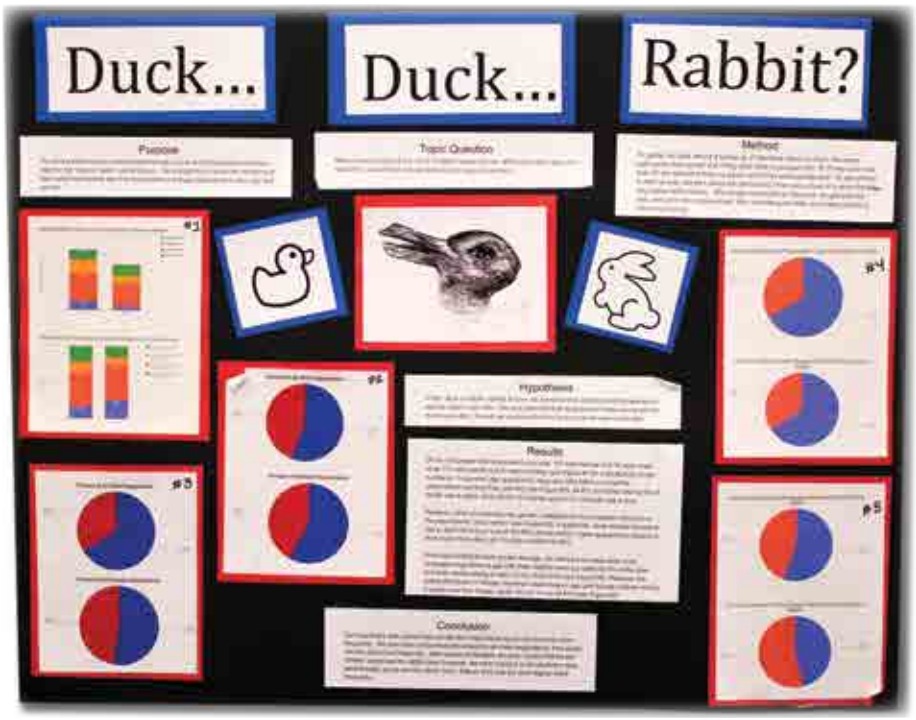


Noah Hellerslia



Owen Hall

Vaishnavi Bhattaram



Oscar Hoffman and Peter Cleary

FIRST PLACE

Noah Hellerslia
"First-Grade Shoelaces"
Highland Elementary School
Abington, Pennsylvania

SECOND PLACE

Owen Hall
"Screen Time"
Odyssey Elementary School
Appleton, Wisconsin

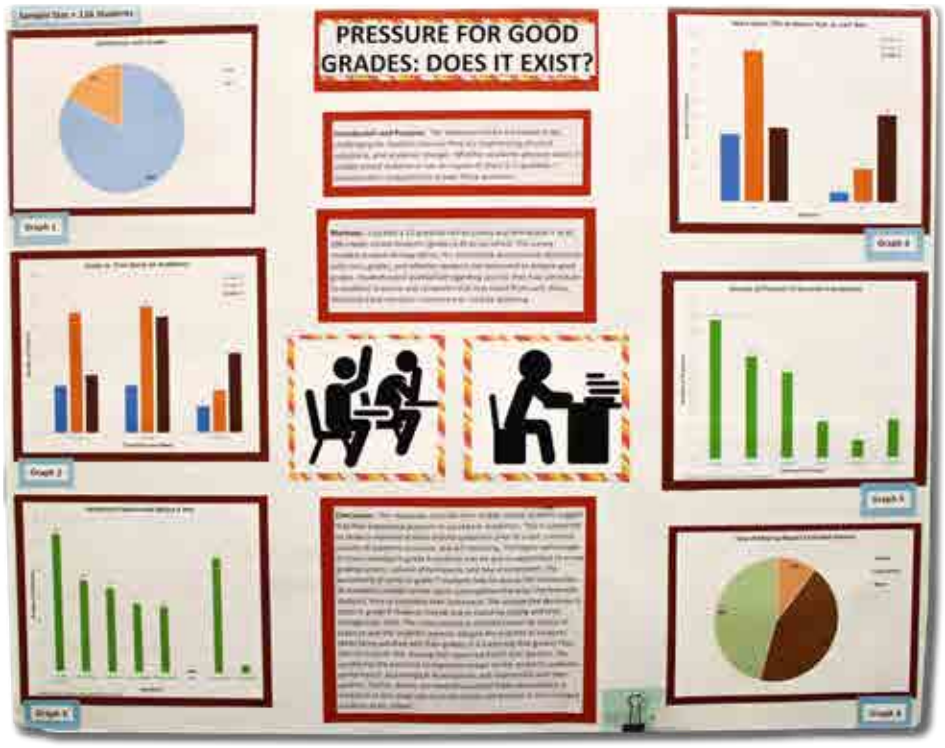
THIRD PLACE

Vaishnavi Bhattaram
"Sugary Soda Showdown"
Jeffers Hill Elementary
Columbia, Maryland

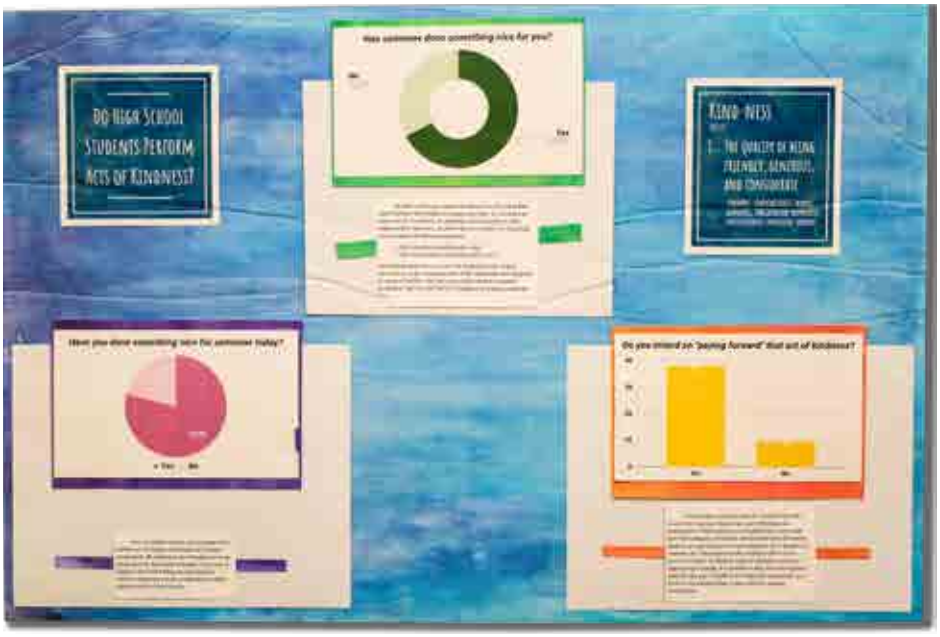
Honorable Mention

Oscar Hoffman and Peter Cleary
"Duck … Duck … Rabbit?"
Rydal Elementary School
Abington, Pennsylvania

POSTER COMPETITION WINNERS GRADES 7–9

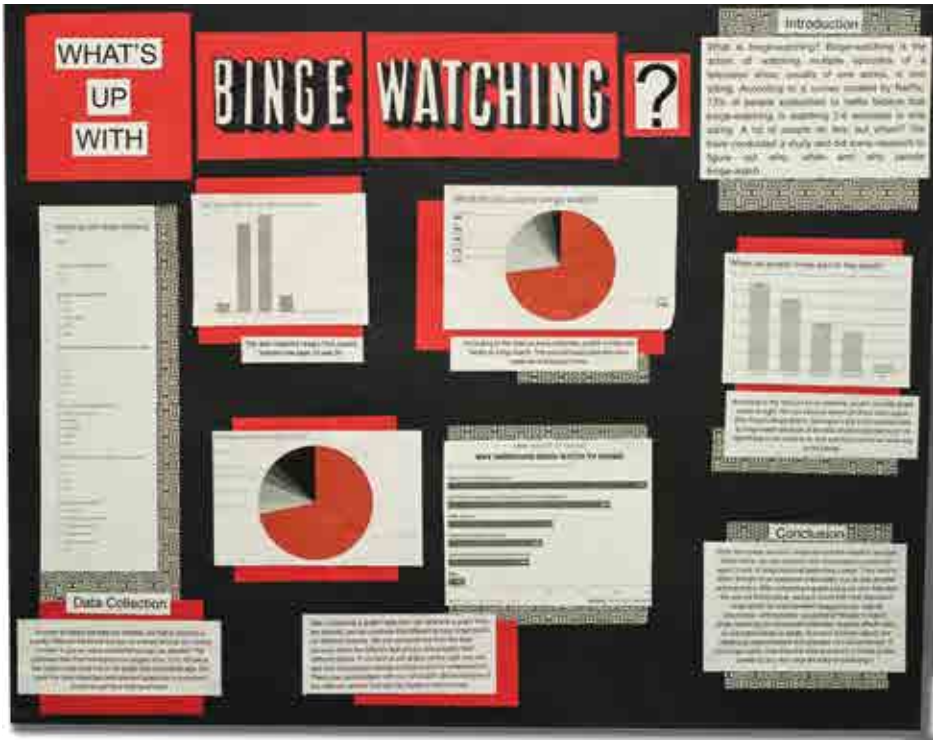


Ella Wong



Esha Pandya and Margaux Castillo

Hannah Wawersik, Anne Marie Singley, Ainsley Lord, and Emma Doherty



Jaclyn Golaszewski, Mickala Winne, and Samantha Dulz

FIRST PLACE

Ella Wong

"Pressure for Good Grades: Does It Exist?"

Hawken School

Lyndhurst, Ohio

SECOND PLACE

Esha Pandya and Margaux Castillo

"Do High-School Students Perform Acts of Kindness?"

Clarkstown South High School

West Nyack, New York

THIRD PLACE

Hannah Wawersik, Anne Marie Singley, Ainsley Lord, and Emma Doherty

"Starbucks Consumption = Mmm!"

Williamsburg Montessori School

Williamsburg, Virginia

Honorable Mention

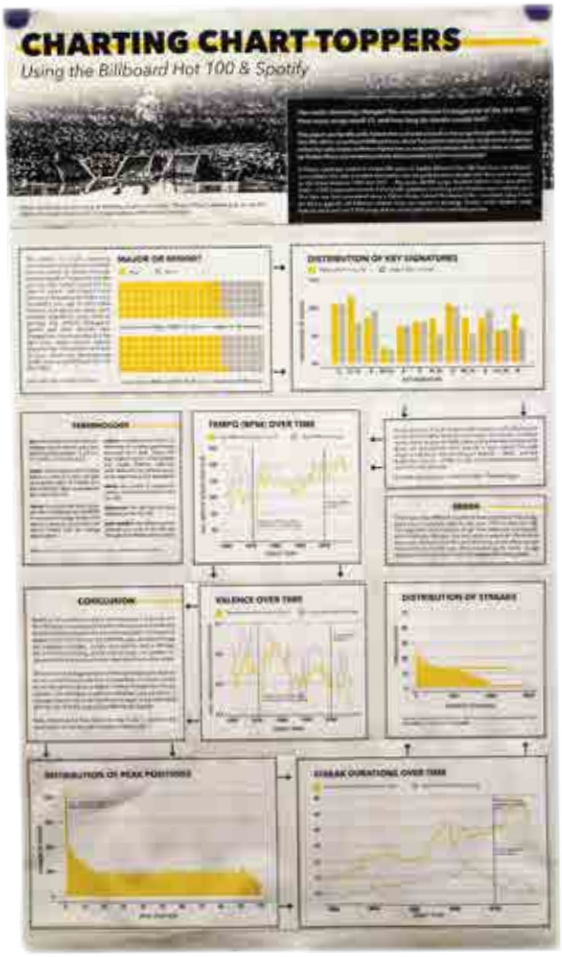
Jaclyn Golaszewski, Mickala Winne, and Samantha Dulz

"What's Up With Binge Watching?"

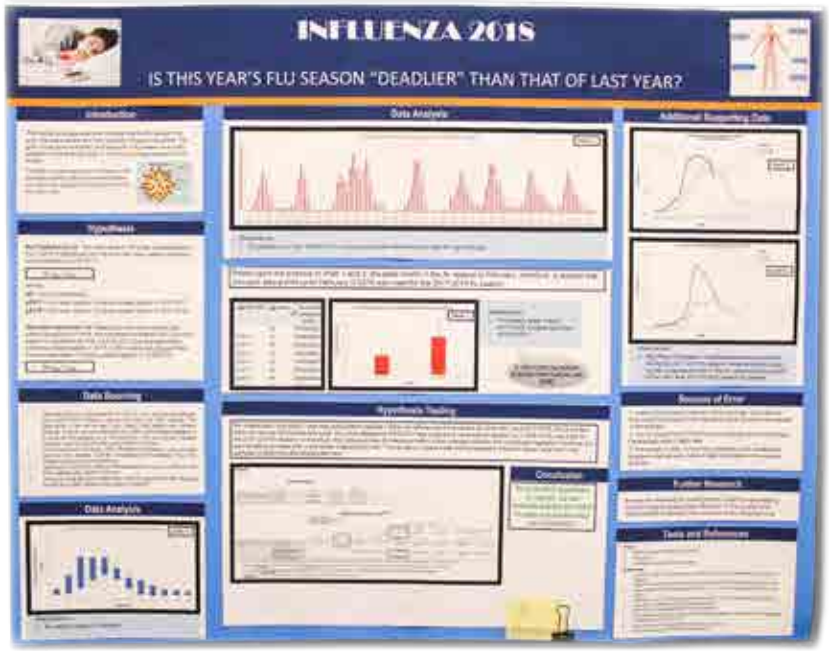
Macomb Academy of Arts and Sciences

Armada, Michigan

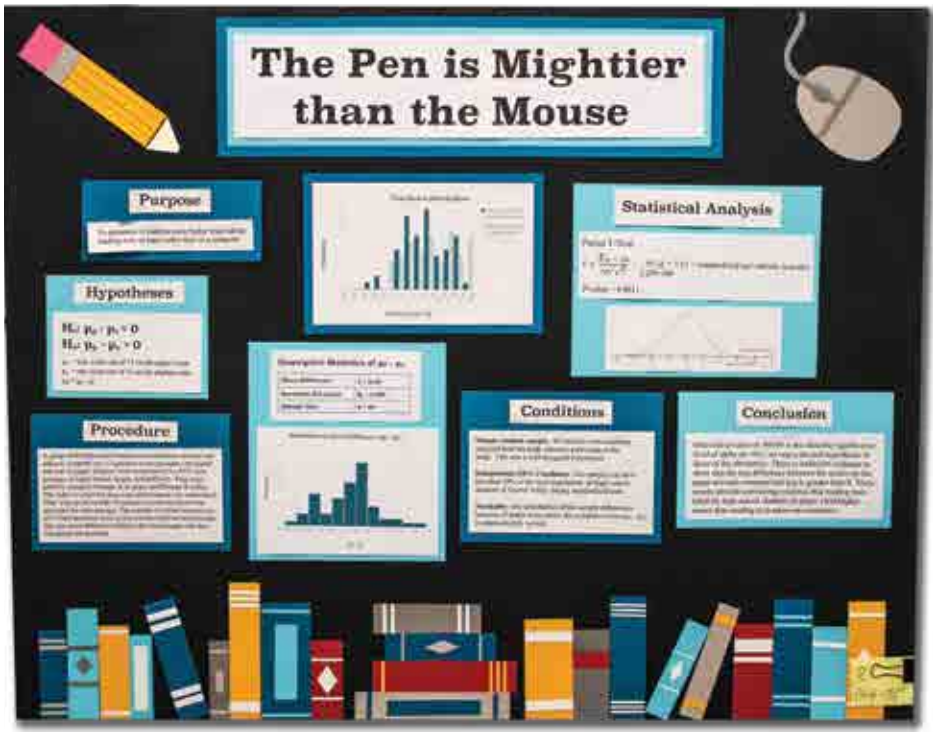
POSTER COMPETITION WINNERS GRADES 10–12



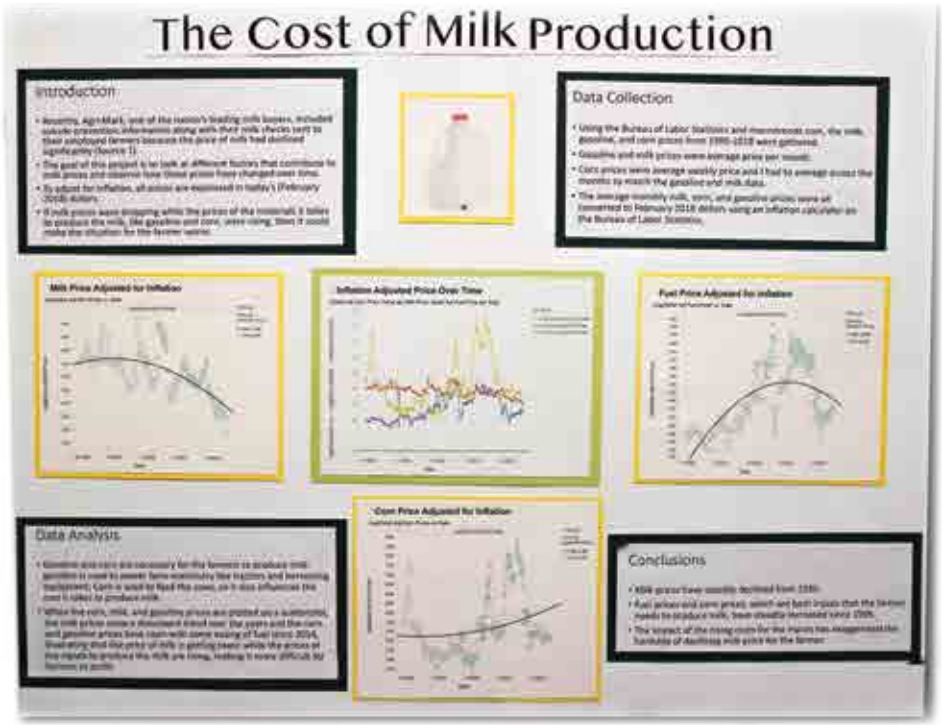
Rudd Fawcett



Dorris Dagama



Dana Caldwell, Ashley Pyon, Sarah Puts, and Lucy Townend



Emma Anziano

FIRST PLACE

Rudd Fawcett

"Charting Chart Toppers"

Phillips Academy

Andover, Massachusetts

SECOND PLACE

Dorris Dagama

"Is This Year's Flu Season Deadlier Than Last Year's?"

Forest Hills Central High School

Grand Rapids, Michigan

THIRD PLACE

Dana Caldwell, Ashley Pyon, Sarah Puts, and Lucy Townend

"The Pen Is Mightier Than the Mouse"

Garnet Valley High School

Glen Mills, Pennsylvania

Honorable Mention

Emma Anziano

"The Cost of Milk Production"

Haddam Killingworth High School

Higganum, Connecticut

The second Akaike Memorial Lecture Award will be presented to **Mike West** of Duke University during the plenary session of the Japanese Joint Statistical Meeting 2018, which will take place at the Korakuen Campus of Chuo University on September 10.

The award, jointly created by the Institute of Statistical Mathematics (ISM) and Japan Statistical Society (JSS), aims to encourage the education of talented young researchers by recognizing those who have achieved outstanding accomplishments that contribute to the field of statistical sciences. It celebrates the outstanding achievements of the late Hirotugu Akaike.

West's contributions to Bayesian statistics include seminal work in dynamic modeling and the implementation of nonparametric models that paved the way for practical data analyses via the first realization of large-scale simulation-based methods. West has also worked at the frontier of various research fields to which Bayesian statistics can be applied and contributed to the creation of data-driven sciences. For example, he established a new approach for biomarker discovery using gene expression data, thus creating a novel trend in -omics biology based on data analysis.

To read more about West's work and the Akaike award, visit *https://bit.ly/2Lge19e.* ■

Noel Cressie, distinguished professor at the University of Wollongong, was recently named a Fellow of the Australian Academy of Science. He was inducted into the academy May 22 in Canberra, Australia.

The academy's citation associated with his induction reads:

"Noel Cressie is a world leader in statistical methodology for analyzing spatial and spatio-temporal data, and its applications to environmental science. His fundamental contributions changed the basic paradigm for analyzing observations in space and space-time. Cressie has also contributed to research on pollution monitoring, climate prediction, ocean health, soil chemistry, and glacier movement and is a NASA science team member for the Orbiting Carbon Observatory-2 mission. Responding to the huge volumes of complex data in environmental research, Cressie has made ground-breaking innovations for 'big data analytics' for remote sensing and climate change."

Read more about the Australian Academy of Science's election of fellows at *https://bit.ly/2LdFGaF.* ■

More than 150 students from across Mumbai came together to attend a two-day statistics workshop, Sankhyiki, in January. The theme was "Data Science and Python Training." In its eighth year, the annual event surpassed all benchmarks.

Sankhyiki featured speakers such as **C.G. Venkatesh**, CoP lead – data science at Larsen and Toubro Infotech; **Madhumita Ghosh**, big data analyst at IBM India Private Limited; **Leslie Rebello**, director of L.R. Associates Private Limited; and **Harshil Gandhi**, associate technical consultant working with SAS Institute India.

The chief guest was **Vyom Upadhyay**, head of analytics and business intelligence at ICICI Bank. He focused on the emerging field of big data, emphasizing that statistics should be used to promote business growth. When asked to pick between proprietary and open source software, he picked open source, reasoning that many minds working together on any idea is better than a few.

Several students were honored with the following awards:

- *Best Summer Internship Award* **Rajitha Nair** and **Hiteshree Lad** (first prize), **Priyanka Joglekar**, and **Cheryl D'Souza**
- *Business Analytics and Data Mining Championship* **Shreya Kekre** and **Shrea Piplani**
- *ISPS Master's of Science Student Project Competition* **Meghna Dineshan**, **Rutika Vagal**, and **Karishma Iyer** (second place)

Taking into consideration the significance of machine learning in the global arena, the second day was dedicated to Python training. Harshil Gandhi, an associate technical consultant working with SAS Institute India, led the five-hour session, which was attended by 60 students.

Sankhyiki 2018 ended by meeting its goal: unveiling the real power of statistics. ■

Four ASU Judges at poster judging, from left: El-Ham Ismail (ASA student member and winner in recent ASA DataFest at ASU), Jonathan Kurka (ASU instructor), Jennifer Broatch (ASU West Campus assistant professor), and Mickey Mandencido (ASU West Campus assistant professor)

Arizona Chapter Hosts Judging for ASA National Data Visualization Poster Competition for K–12

The Arizona Chapter concluded its academic year of activities by supplying judges for the ASA's Data Visualization Poster Competition for K–12 students. Hosted by Jennifer Broatch at Arizona State University West Campus, the final national round of judging included 134 posters submitted from the regional competitions. The judges came from both ASU faculty and statistics students, as well local industry—18 altogether, including coordinator and chapter president Rodney Jee.

The event not only served to determine the national winners for this annual competition, but also provided an opportunity for students, professors, and industry professionals to exchange views about graphics and statistical ideas.

Using a recently revised rubric to award the posters on their application of data visualization, the judges often found posters with impressive work (occasionally beyond the AP curriculum) and/or amusing subjects.

Judging for the nationals rotates throughout the ASA's chapters. It is very much an "event in a box," since the regional competitions mail the posters to the national judging site and about a dozen chapter members spend a full day judging before communicating the results to the ASA. The ASA provides the awards to the national winners, coordinates the regional sites with the national judging coordinator, and supports a lunch for the judges. ■

See **Page 30** for all the poster winners for 2018.

From left: Bob Peterson, Karry Roberts, Kathy Peterson, Jaclyn Golaszewski, Mickala Winne, Samantha Dulz, Rob Kushler, and Violet Fiddes. The three poster winners won an honorable mention in the Grades 7-9 category of the competition for their poster titled, "What's Up with Binge Watching?"

Detroit, Ann Arbor Chapters Recognize K–12 Poster Winners

Karry Roberts, ASA Detroit Chapter Secretary

Members of the ASA Detroit and Ann Arbor chapters recently presented awards for the ASA National Data Visualization Poster Competition and Michigan Statistics Poster Competition to students at their schools.

For several years, Robert Kushler, a Detroit Chapter board member, has coordinated award presentations for students in southeast Michigan with sponsors of the Michigan event currently led by Dan Adrian at Grand Valley State University (GVSU). Offered as an incentive to sponsoring teachers, members of the Detroit Chapter travel to the schools for a ceremony, rather than just mailing the plaques and certificates. They also present the teachers with a statistical book in recognition of their contributions.

See **Page 30** for all the poster winners for 2018.

This year, the chapter held two award presentations. A team of ninth-grade students from Macomb Academy of Arts and Sciences in Armada, Michigan, received an honorable mention at the national level. Violet Fiddes, a mathematics and computer science teacher there, encouraged her students to enter the competition.

The poster by Samantha Dulz, Jaclyn Golaszewski, and Mickala Winne, titled "What's Up with Binge Watching?" won an honorable mention award in the Grades 7–9 category of the National Data Visualization Poster Competition. They also won a second-place award at the state-level Michigan Statistical Poster Competition.

Also, in Fiddes' class, "Printed vs. Ebooks" by Kayla Whitney and Jessica Jarema received a third-place award and "Do You Know Crime" by Cameron Keller, Jacob Brown, and Brennan McClelland was a national qualifier at the Michigan Statistics Poster Competition.

From left: Nicholas Moloci, Paavani Tewari, Nicole Pike, Rob Kushler, Anamaria Kazanis, and Karry Roberts at Uriah Lawton Elementary School, Ann Arbor, Michigan. Third-grader Tewari won first place in the Michigan Statistics Poster Competition.

Representing the ASA Detroit Chapter Board for this event were Rob Kushler, Kathy Peterson, Bob Peterson, and Karry Roberts.

During another recognition event at Uriah Lawton Elementary School in Ann Arbor, Anamaria Kazanis, council of chapters representative to the ASA Board; Nicholas Moloci, vice president of the Ann Arbor Chapter; and Kushler and Roberts from the Detroit Chapter presented a first-place Michigan Statistics Poster Competition award to third-grade student Paavani Tewari for "Exploring Public Library Summer Game Data." ■

sectionnews

Physical and Engineering Sciences

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

- What an applied statistician does
- How an applied statistician solves problems in science, engineering, technology, and business
- What nontechnical skills are required to be successful as an applied statistician

The Marquardt Industrial Speakers Program was established by SPES in the early 1990s to encourage careers in applied statistics. If you are an institution interested in having a speaker or a SPES member interested in being on the speakers list (or working directly with a local institution to set up a visit), contact Vaneeta Grover at *vkgrover@yahoo.com.* ■

Biopharmaceutical

The ASA-BIOP Nonclinical Biostatistics Working Group (NCBWG) endorsed the creation of a new workstream co-chaired by Phillip Yates of Pfizer Inc. and Katja Remlinger of GSK. The goal of the workstream is to provide information, insight, and networking opportunities for students interested in careers in the nonclinical area and offer a forum for young professionals to more actively engage in ASA-BIOP activities.

In addition to supporting the Nonclinical Biostatistics Conference, we are interested in pairing industry veterans with nearby academic partners. Contact Yates (*phillip.yates@pfizer.com*) or Remlinger (*katja.s.remlinger@gsk.com*) or visit the NCBWG website at *https://bit.ly/2Nmyt9h* for more information or to volunteer. ■

Stan Sclove of the University of Illinois at Chicago tells old funny stories at the Bernard Harris Memorial Symposium, which honored the first chair of the ASA Section on Risk Analysis. The symposium was held in Raleigh, NC, in May.

People, Ideas Come Together at Risk Analysis Symposium

Alexandra Kapatou

Most conferences help people connect with colleagues from distant places. A few conferences help those who attend learn new things. It is a rare conference that creates an "Aha!" moment, where people and ideas come together. The Bernard Harris Memorial Symposium was this rare type of conference.

The idea to organize a symposium in honor of Bernard Harris, the first chair of the ASA Section on Risk Analysis, arose a couple of years ago when Susan Harris, Bernard's widow, presented the section with a gift. The gift indicated it was meant to cover expenses of invited speakers from interdisciplinary areas, like Bernard. What is risk analysis, if not an interdisciplinary area, after all? We imagined an intimate, small conference, where a few invited speakers would have plenty of time to present their research and the audience would have enough time to ask questions and give feedback. It took place May 10–11 in Raleigh, North Carolina.

The speakers, in order of presentation, included the following:

- **Michael Pennell** of The Ohio State University and chair of the Risk Analysis Section gave the opening remarks.

Clarice Weinberg and David Banks discuss risk at the Bernard Harris Memorial Symposium.

- **Stan Sclove** of the University of Illinois at Chicago told us about the life and work of Bernard Harris.
- **Edward Melnick** from NYU Stern School of Business began the technical part of the symposium by presenting the foundations of risk assessment.
- **Richard Smith,** from The University of North Carolina at Chapel Hill and associate director of the Statistical and Applied Mathematical Sciences Institute (SAMSI), presented his research on risk of extreme weather events in a changing climate.
- **R. Dale Hall** from the Society of Actuaries presented segmentation and decomposition techniques in actuarial risk analysis using predictive analytics.

R. Dale Hall from the Society of Actuaries presents actuarial risk analysis techniques using predictive analytics.

From left, front row: Edward Melnick, Stan Sclove, David Banks, and Susan Harris; back row: Richard Smith, Clarice Weinberg, Erin Charles, and John Wambaugh

Richard Smith, left, discusses extreme weather events. Seated are Aleka Kapatou, Michael Pennell, and Clarice Weinberg.

- **John Wambaugh** of the US Environmental Protection Agency presented informatics tools for chemical safety.

The talks were followed with poster presentations by graduate students Taeho Kim and Shiwen Shen from the University of South Carolina and Jun Shepard from Duke University. The first day of the symposium closed with a mixer.

On the second day, **Clarice Weinberg** from the Biostatistics and Computational Biology Branch of the National Institute of Environmental Health Sciences, answered the question: "Can we develop a way to find the multi-SNP contributors to disease risk?" **David Banks**, from Duke University and director of SAMSI, presented methods of adversarial risk analysis. **Ilyan Iliev** of the University of Southern Mississippi discussed terrorism and social media. **Rita Fuller** from New York Life Insurance Company gave a talk for the benefit of students on preparing for a job interview in data science. Finally, **Matthew Wheeler**, from the National Institute for Occupational Safety and Health and past chair of the Risk Analysis Section, gave closing remarks by presenting inspired ideas that put together different areas of risk analysis. ■

Symposium Organizing Committee Members

Michael Pennell, The Ohio State University

Matthew Wheeler, National Institute of Occupational Safety and Health

Susan Simmons, North Carolina State University

Alexandra Kapatou, American University

Maria Barouti, American University

Qian Li, FDA Center for Tobacco Products

Piaomu Liu, Bentley University

Mary Louie, New York Life Insurance

Edsel Peña, University of South Carolina

Chris Sroka, New Mexico State University

Wensong Wu, Florida International University

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.

For display and online advertising rates, go to *www.amstat.org/ads*.

For questions email *advertise@amstat.org*.

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

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

Connecticut

■ Join Boehringer Ingelheim's team as a Sr. Principal Methodology Statistician in Ridgefield, CT! The Sr. Principal Methodology Statistician acts as an international specialist in applied statistical methodology to foster methodological innovation in Boehringer Ingelheim (BI) Clinical Statistics and develop innovative, efficient statistical methods for ready-to-use application in BI studies and projects. Use the following link to submit your resume for this opportunity: *https://bit.ly/2NRPCZk* EOE.

Massachusetts

■ The Department of Biostatistics and Computational Biology (BCB) at the Dana-Farber Cancer Institute seeks an experienced and highly motivated PhD biostatistician to engage collaboratively with investigators on basic science, animal model, and human research activities in multiple areas of adult

Westat | An Employee-Owned Research Corporation®

EOE

Statistical Career Opportunities with Westat

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

We are currently recruiting for the following position:

Survey Sampling Statistician

This position requires a master's degree or Ph.D. in statistics with coursework in survey sampling or a master's or Ph.D. in survey sampling. Candidates with a master's must have at least 8 years of experience in sample survey design, selection, or weighting and a Ph.D. with 6 years' experience. Although not required to do programming, candidates would benefit from knowing SAS and other statistical software packages. Qualified candidates must have excellent written and oral communication skills, strong organizational skills, and the ability to handle multiple tasks simultaneously.

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

www.westat.com

The Williams College Department of Mathematics and Statistics invites applications for a new tenure-track position in Statistics, beginning fall 2019, at the rank of assistant professor. A more senior appointment is also possible for a qualified candidate at a later stage in their career. The candidate should have a Ph.D. in Statistics or a closely related field by the time of appointment. We are seeking candidates who show evidence and/or promise of excellence in teaching and a strong research program that can engage undergraduate students. The candidate will become the seventh tenure-track statistician in the department, joining a vibrant and innovative group of statisticians with an established statistics major. For more information on the Department of Mathematics and Statistics, visit *http://math.williams.edu/*.

Candidates may apply via *https://apply.interfolio.com/50978* by uploading a cover letter addressed to Professor Richard De Veaux, a curriculum vitae, a teaching statement, a description of research plans, and three letters of recommendation on teaching and research. The Department is committed to building a diverse and inclusive community. In your application materials, we also ask you to address how your teaching, scholarship, mentorship and/or community service might support Williams's commitment to diversity and inclusion.

Expectations: The teaching load is two courses per 12-week semester and a winter term course every other January. The candidate will be expected to teach introductory statistics, core courses for the statistics major, and elective courses in their areas of interest. The successful candidate will establish an independent research program that results in scholarly publications. Williams College provides broad support for start-up funds, funding for student research assistants, faculty professional development funds, and a shared computer cluster for parallel computation.

Review of applications will begin on or after October 1st and will continue until the position is filled. All offers of employment are contingent upon completion of a background check. Further information is available at *https://faculty.williams.edu/prospective-faculty/background-check-policy/*. Williams College is a coeducational liberal arts institution located in the Berkshire Hills of western Massachusetts. The college has built its reputation on outstanding teaching and scholarship and on the academic excellence of its approximately 2,000 students. Please visit the Williams College website (*http://www.williams.edu*). Beyond meeting fully its legal obligations for non-discrimination, Williams College is committed to building a diverse and inclusive community where members from all backgrounds can live, learn, and thrive.

Williams College

oncology and HIV disease. PhD and at least 2 years of collaborative experience are required. Prior experience in oncology and/or HIV is a plus. *https://bit.ly/2mo36j8* EOE.

Texas

■ The Department of Mathematics and Statistics at Texas Tech University announces a tenure-track position opening in applied statistics (biostatistics, statistical genetics, data science) – requisition #14433BR. Candidates need to apply at *https://bit.ly/2JqpFwA.* In addition to the mandated online application, applicants must have three reference letters emailed to *misty.rangel@ttu.edu*. Review of applications will begin Sept. 1, 2018. Applications will be accepted until Dec. 1, 2018. EOE. ■

AMSTATNEWS

ADVERTISING DIRECTORY

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

misc. products and services

professional opportunities

software

SOCIAL CHATTER

What our followers are saying online

FOLLOW US

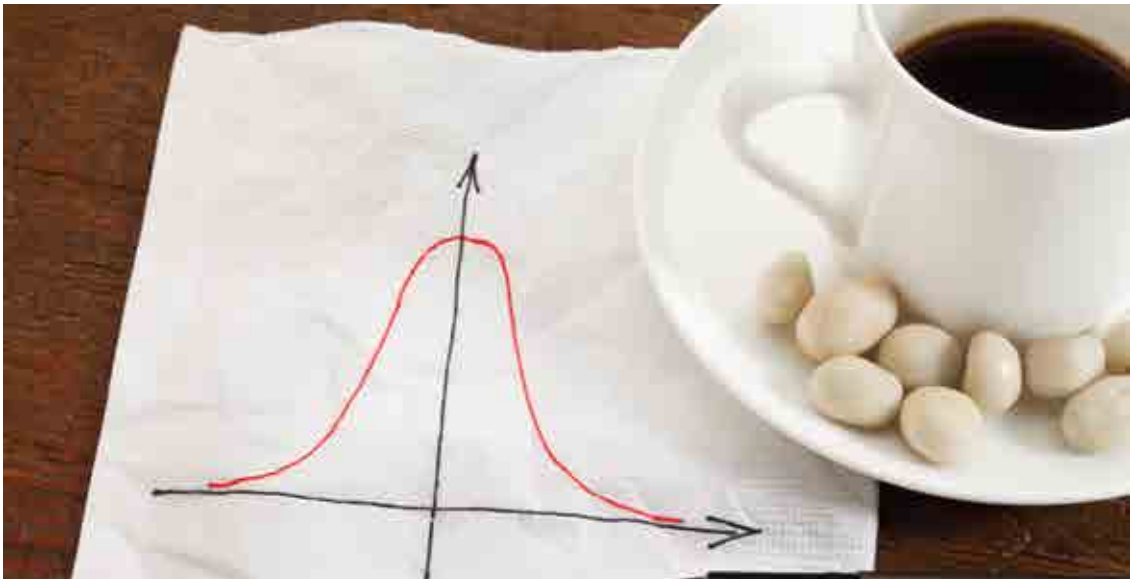
community.amstat.org

www.facebook.com/AmstatNews

@AmstatNews

www.instagram.com/AmstatNews

What is one thing a statistician cannot live without?



Next month, we'll ask: If a statistician were a Disney character, who would it be? Tag @AmstatNews with your answer.

Morné Lamont The normal distribution :-)

Justin Smith Long-term relative frequency definition of probability, strong law of large numbers, central limit theorem

Mohammed Qurish I think two; certainly the variability and principles.

Simon Lessard Bonaventure We are like garbage workers; our life is meaningless without residuals

Ronnie Pingel A clearly defined aim of the study

Kshitij Kulkarni *P* value

Arturo Gerardo Explaining important math results to non-statisticians.

Aidan Bindoff
@ABindoff
Coffee

VlakePhoenix
@VlakePhoenix
The ability to quantify risk.

Michelle
@StatsN3rdGirl
Ethics

statverse
@statverse
Exciting collaborations

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.