The Department of Statistics and Data Sciences is an academic unit housed in the College of Natural Sciences that supports the statistical and data science needs of The University of Texas at Austin campus.

**Our Mission**

Our mission is to be a world-class center for statistical science that advances scholarship across The University of Texas at Austin and prepares graduates to flourish in an information-rich world.

**Our Vision**

The Department of Statistics and Data Sciences will become a vibrant community of scholars that has a strong intellectual identity and is widely recognized for excellence in interdisciplinary research and high impact educational programs.

**Our Core Values**

- Commitment to education
- Excellence in research
- Broad engagement with science and society
- Collegiality and integrity
From the Chair

It is my pleasure to present the Department of Statistics & Data Sciences Annual Report on the activities during academic year 2016–17. There is much to be proud of in this report as many of the achievements and accomplishments of our faculty, students, and staff illustrate the impact of our collective efforts to fulfill our mission to be a world-class center for statistical science that advances scholarship across The University of Texas at Austin and prepares graduates to flourish in an information-rich world. I would like to share some highlights from this past year:

10th Annual Summer Statistics Institute: in May 2017, we hosted the 10th Annual Summer Statistics Institute (SSI). Participants could select from 23 12-hour courses spanning topics from introductory statistics, statistical software, statistical methods and statistics applications. Over the last decade, SSI has welcomed a total of 5546 participants, relied on 258 UT graduate student volunteers, and employed 58 instructors (including six who have taught for each of the ten years).

First PhD students graduating: in May 2017 Oscar Hernan Madrid Padilla was the first student to earn a PhD in Statistics from The University of Texas at Austin. Following closely is Tianjian Zhou who defended his dissertation in June and will graduate in August. Both have accepted postdoctoral positions: Padilla at the University of California–Berkeley and Zhou at NorthShore University HealthSystem Research Institute.

Dr. Kristin Harvey was recognized for her exemplary work teaching freshmen with a well-deserved Dads’ Association Centennial Fellowship Award.

Other notables include:

— the initial cohort of trainees for the NIH Biomedical Big Data training grant successfully completed their first year in the program. Trainees include Vikram Baruah (PhD student in Biomedical Engineering), Evan Ott (PhD student in Statistics), Raghav Shroff (PhD student in Cellular & Molecular Biology), Kevin Song (PhD student in Computer Science), and Serena Zhao (PhD student in Ecology, Evolution, & Behavior).

— successful faculty recruitment of Dr. Abhra Sarkar. SDS welcomes Dr. Abhra Sarkar in fall 2017. Dr. Sarkar earned his PhD in Statistics from Texas A & M and joins us directly from a postdoctoral fellowship at Duke University. Also, Dr. Jared Murray will join the SDS core faculty. Dr. Murray was hired by the IROM Department and arrives most recently from the Department of Statistics at Carnegie Mellon University.

— Dr. Kam Hamidieh joined the SDS, taking on the role of expanding our industry connections and overseeing the Graduate Fellows program.

— the award of new funded research projects, including Dr. Purna Sarkar’s award from the National Science Foundation for her project, *Inference for network models with covariates: Leveraging local information for statistically*.

— my work as a panel member for the Review of the Compliance, Safety, and Accountability (CSA) program of the Federal Motor Carrier Administration (FMCSA). Sponsored by the National Academies of Sciences, Engineering, and Medicine the panel developed recommendations to modify the motor carrier evaluation system and collect better data for safety. Details about the report: [https://www.eurekalert.org/pub_releases/2017-06/naos-nrf062717.php](https://www.eurekalert.org/pub_releases/2017-06/naos-nrf062717.php)

Plus, we continue exploring opportunities for further growth by building partnerships with the Dell Medical School on faculty hires, consulting, and training.
Looking forward, we are hosting the 2017 International Workshop on Objective Bayes Methodology this December. O’Bayes 2017 is the 12th meeting in a series of workshops held to discuss recent developments in Bayesian methods for inference in increasingly larger scale scientific research problems.

And as a sidenote, I will be stepping down as chair in August and returning to my previous institution, the University of Florida. I have thoroughly enjoyed my time at UT and in Austin and am extremely excited about the future of SDS. I will be following the growth closely from Florida.

I would also like to take this opportunity to thank the dedicated faculty and staff of the department for working with me in establishing the department and also President Gregory Fenves, Executive Vice President and Provost Maurie McInnis and CNS Dean Linda Hicke for their continued support.

Thank you for another great year.

Dr. Mike Daniels
**Table of Contents**

I. SDS Highlights in 2016–17 ............................................................................................................. 1

II. Unit Reports

A. Instruction ...................................................................................................................................... 6
B. Consulting ....................................................................................................................................... 10
C. Graduate Degree Programs ........................................................................................................... 12
D. Graduate Fellows Program ............................................................................................................. 14
E. Portfolio and Certificate Programs ................................................................................................. 15
F. SDS Seminar Series ....................................................................................................................... 18
G. Summer Statistics Institute ............................................................................................................. 19
H. Corporate Sponsorships .................................................................................................................. 20
I. Grants .............................................................................................................................................. 21
J. Development ................................................................................................................................... 23

III. Appendices

A. Organizational Chart ...................................................................................................................... 24
B. SDS Core Faculty ........................................................................................................................... 25
C. SDS Seminar Series Speakers ...................................................................................................... 30
I. HIGHLIGHTS IN 2016–17

BIG DATA TO KNOWLEDGE (BD2K): T32 Predoctoral Training in Biomedical Big Data

In 2016 Dr. Michael Daniels was awarded a five-year T32 (NIH) predoctoral training grant in biomedical big data for a total project amount of $1,018,560. The BD2K project was developed to deepen predoctoral students’ knowledge in three key areas: biology, statistics, and computer science in order to provide the holistic training necessary to meaningfully mine biomedical big data.

In Fall 2016 the Executive Committee selected the first cohort of trainees:
- Vikram Baruah (PhD student in Biomedical Engineering)
- Evan Ott (PhD student in Statistics)
- Raghav Shroff (PhD student in Cellular and Molecular Biology)
- Kevin Song (PhD student in Computer Science)
- Serena Zhang (PhD student in Ecology, Evolution, and Behavior)

The trainees began their training by taking the three core courses and the seminar/workshop course, as well as attending the NIH BD2K consortium meeting in Bethesda, Maryland in November 2016. Core courses include SDS 385: Statistical Models for Big Data, CSE 380: Tools and Techniques for Computer Science and BIO 382K: Introduction to Biology for Data Science. Both SDS 385 and BIO 382K were specifically developed for the training program by SDS/IROM faculty member Dr. James Scott and by Biology faculty member Dr. Hans Hoffman respectively.

In this next project year, three of the original five trainees will each do two rotations in areas of their choosing, receiving direct mentoring from faculty working with biomedical big data. The other two trainees will complete up to one rotation and focus on their doctoral research. Two new trainees have been selected to start the program in Fall 2017.

10th Annual UT Summer Statistics Institute

The Department of Statistics and Data Sciences (SDS) hosted the 10th annual UT Summer Statistics Institute (SSI) May 22–25, 2017, in collaboration with Academic Technology Support in the College of Liberal Arts building. Participants could select from 23 twelve-hour courses designed to appeal to a broad range of students, faculty, staff, and the general public. Participants joined the Institute from across the country, coming from as far away as Vietnam, illustrating the reach and popularity of SSI worldwide.

SSI offers intensive four-day workshops on diverse topics from introductory data sciences to advanced statistics. Whether participants are new to data analysis or seasoned statisticians, SSI provides a unique hands-on opportunity to acquire valuable skills directly from experts in the field. Through the last decade, SSI has welcomed a total of 58 instructors, 258 UT graduate student volunteers, and 5546 participants. This year, Camry Isaac, a sophomore in the University Leadership Network, joined SDS as the SSI coordinator-in-training under the guidance of Sasha Schellenberg.

SDS is pleased to recognize six instructors who have participated in SSI for each of the last 10 years:
- Dr. Stephen Jessee, Department of Government
- Dr. Michael Mahometa, Department of Statistics and Data Sciences
- Dr. Marc Musick, Department of Sociology
• Dr. Keenan Pituch, Department of Educational Psychology
• Dr. Tom Sager, Department of Information, Risk, and Operations Management
• Dr. Tiffany Whittaker, Department of Educational Psychology

Special recognition goes to the 2017 Ice Cream Network Sponsors: CaboBobs, Moojo, Alamo Draft House, Amy’s Ice Cream, and P Terry’s Burger Stand.

**PhD in Statistics Program Celebrates its First Doctoral Graduates**
The PhD in Statistics program celebrates its first two graduates of the program: Dr. Oscar Madrid Padilla and Dr. Tianjian Zhou. Oscar wrote his dissertation, *Constrained Estimation via the Fused Lasso and Some Generalizations* under the supervision of Dr. James G. Scott. In July 2017, he will start his postdoctoral fellowship at the University of California–Berkeley working with Dr. Michael Jordan. Tianjian wrote his dissertation, *PairCloneTree: Reconstruction of Tumor Subclone Phylogeny Based on Mutation Pairs using NGS Data*, under the supervision of Dr. Peter Mueller. Tianjian will start his postdoctoral fellowship at the NorthShore University HealthSystem, Chicago, in July 2017.

Both graduates were part of the inaugural PhD cohort that started the program in Fall 2013. The PhD in Statistics program was approved by the Texas Higher Education Coordinating Board in November 2012 and provides students with rigorous training in advanced statistical methods and first-hand experience applying and developing such methods to analyze complex data in diverse fields.

**Faculty Search**
The Department of Statistics & Data Sciences (SDS) conducted a year-long search for a new junior faculty position that sits 100% in SDS. After reviewing 118 applications, six applicants were invited for on-campus interviews. A faculty position was offered to one candidate who will be joining SDS in Fall 2017:

- **Abhra Sarkar** comes to UT from the Department of Statistical Science at Duke University. Dr. Sarkar received his PhD in Statistics from Texas A&M University. He will be appointed as an Assistant Professor in SDS. His main research focus is on Bayesian semiparametric and nonparametric methods in the context of latent variable models, Markov models, and measurement error models.

**New Staff**
The Department of Statistics and Data Sciences welcomed three new staff members to the administrative team this year, Stephanie Tomlinson (Grants and Contracts Specialist), Ramzi Saud (Student Employee) and Camry Isaac (ULN Intern).

- **Stephanie Tomlinson** comes to SDS from Biomedical Engineering where she worked as an Administrative Associate. Stephanie was recognized in 2017 for her 15 years of service with The University of Texas at Austin.
- **Ramzi Saud** comes to SDS from The University of Texas Information Technology department where he previously held the position of Student Technician.
- **Camry Isaac** comes to SDS from the University Leadership Network program to assist in the planning, implementation, and logistical support of the 10th Annual UT Summer Statistics Institute.
Faculty Awards
Dr. Kristin Harvey was recognized for her outstanding contributions teaching freshmen by winning the Dads’ Association Centennial Teaching Fellowship award. This fellowship was established in 1983 to recognize faculty members who are actively engaged in the instruction of freshman undergraduates. Dr. Harvey teaches and coordinates SDS 302: Data Analysis for the Health Sciences and has been with the department as a lecturer since August 2013.

Dr. Tom Sager won the best paper prize for “The Risk of Variable Annuity Guarantees and Life Insurer Capital” (with Professors Etti Baranoff and Bo Shi), Asia-Pacific Journal of Risk and Insurance, 2016. The $1,000 award was presented at the annual meeting in Chengdu, China.

Visiting Researchers
The Department of Statistics and Data Sciences hosted three visiting researchers this year, Julyan Arbel, Lorenzo Capello, and Iraj Kazemi.

- **Julyan Arbel** comes from Inria Grenoble – Rhône-Alpes, France, where he is a member of the Mistis team. He earned his PhD in Statistics from Crest – Université Paris-Dauphine in 2013. His research interests are in Bayesian (parametric and nonparametric) statistics with applications in environmental science, biodiversity, and species models. During his residency at UT he worked with Dr. Stephen Walker.

- **Lorenzo Capello** comes from Bocconi University in Milan, Italy, where he is a PhD student candidate. He received his bachelor’s degree in Naval Engineering with a focus on applied mathematics and physics. He then continued his studies with a master’s in Statistics and Finance, where he became passionate about time-series modeling. His areas of interest are Bayesian nonparametric inference, nonparametric regression, Markov models, and time-series analysis. During his residency at UT he worked with Dr. Stephen Walker.

- **Iraj Kazemi** comes from the University of Isfahan, Iran, where he is an Associate Professor of Statistics. He earned both his bachelor’s and master’s degrees in Statistics from the Shiraz University, Iran, and his PhD in Applied Statistics from Lancaster University, UK. His areas of interest are generalized linear mixed-effects models, semiparametric Bayesian methods, and flexible modeling of random effects distributions in longitudinal data. During his residency at UT he worked with Professor Mike Daniels.

New Course: SDS 301 Elementary Statistical Methods
SDS has received approval in Fall 2016 to offer a new course: SDS 301 Elementary Statistical Methods. This course is designed to replace M 316 Elementary Statistical Methods that was retired in Spring 2015. SDS 301 fulfills both the core math component as well as the quantitative reasoning flag. Additionally, students will be able to apply for transfer credit for statistics courses taken at community colleges or AP Statistics credit. Dr. Kristin Harvey is developing the course materials for the first course offering in Fall 2017. The class will begin with one section of 96 students, with plans to offer additional sections in the future.

The course will have a strong focus on statistical reasoning with the course objectives of:

1) learn the basics of research design and sampling procedures.

2) learn the basics of data analysis, including the elementary statistical methods commonly used in a variety of fields.

3) learn how to analyze and model real-world data using common software programs.

The creation of this course allows SDS to better support all students in the university to ensure they are enrolled in a statistics class which meets their degree plan needs while taking a course.
that focuses on topics most relevant to their field of study. Students enrolled for Fall 2017 represent the following colleges: 62% Liberal Arts, 15% Undergraduate Studies, 7% Communications, and 6% Fine Arts, with Social Work, Natural Sciences and Education comprising the final 10%. The large proportion of Liberal Arts and Fine Arts majors is unique to this class. The current undergraduate classes offered in SDS primarily support students from the College of Natural Sciences and Moody College of Communication.

**Dell Medical School Consultation**

In April 2017, the Department of Statistics and Data Sciences Statistical Consulting Group officially opened a new service specifically for the Dell Medical School. This new service is aimed at helping current students and faculty members of the Dell Medical School. A specific website was created to intake any new requests for consultation:
https://stat.utexas.edu/consulting/dell-medical-school-consulting

**UT Community & Committee Service**

SDS Staff served on eight volunteer committees in 2016–17, in addition to participating in several volunteer-based events that serve the greater College of Natural Sciences (CNS) and UT communities. Staff members served on the CNS, Staff and Faculty Councils as well as a variety of academic advisory councils. Staff members also volunteered for the Teaching Discovery Day at the College of Natural Sciences and the College of Natural Sciences Talent Show.

**2016–17 Committee Service**

<table>
<thead>
<tr>
<th>Staff Member</th>
<th>Committee</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kristin Harvey</td>
<td>21st Century Curriculum Planning Implementation Task Force, College of Natural Sciences</td>
<td>Member</td>
</tr>
<tr>
<td></td>
<td>Faculty Advisory Committee for Texas Success Initiative</td>
<td>Member</td>
</tr>
<tr>
<td></td>
<td>STEM Council</td>
<td>Participant</td>
</tr>
<tr>
<td></td>
<td>Faculty Advisory Committee for Quantitative Reasoning Flag</td>
<td>Member</td>
</tr>
<tr>
<td></td>
<td>Course and Curriculum Committee, College of Natural Sciences</td>
<td>Member</td>
</tr>
<tr>
<td>Vicki Keller</td>
<td>General Faculty Standing Committee: Responsibilities, Rights, and Welfare of Graduate Student Academic Employees Committee</td>
<td>Member</td>
</tr>
<tr>
<td>Rachel Poole</td>
<td>University Staff Council</td>
<td>Member, Chair of the Health &amp; Wellness Committee</td>
</tr>
<tr>
<td>Sasha Schellenberg</td>
<td>CNS Staff Council</td>
<td>Member</td>
</tr>
</tbody>
</table>
University Leadership Network (ULN)
This year, SDS partnered with The University Leadership Network (ULN), a nationally recognized incentive-based scholarship program for students. Each year, 500 freshmen are selected to begin the ULN program that includes a comprehensive four-year plan involving leadership training, experiential learning opportunities, and community and university service.

Camry Isaac, a sophomore in the University Leadership Network, joined SDS as the Summer Statistics Institute (SSI) coordinator-in-training under the guidance of Sasha Schellenberg. Camry took the lead role in the organization of the 2017 Summer Statistics Institute, learning valuable skills in communication, organization and personnel management. Camry will return for the upcoming year, taking on the lead role for SSI 2018 with increased responsibility and duties.

With the success of the department’s first ULN intern, Sasha Schellenberg and Rachel Poole developed a 2-year rotation for all incoming ULN Interns. Interns in the first year of the program will focus on the organization and logistical administration of the SDS Seminar Series and will be responsible for SDS social media. In the second year of the program, the intern will take the lead role in the organization of the Summer Statistics Institute. SDS will welcome its second ULN intern in fall 2017.
II. UNIT REPORTS

A. INSTRUCTION

The demand for statistics and scientific computation courses continues to grow, particularly at the service course-level. SDS has responded by creating new courses and adding additional sections of in-demand courses to meet the needs of the diverse UT student population. SDS now offers more than 100 undergraduate and graduate courses each year, ranging from introductory data analysis and undergraduate honors statistics, to advanced graduate topics in Bayesian modeling and computational statistics.

Enrollment Trends

SDS faculty taught 76 undergraduate courses and 40 graduate courses to a total 6040 students in 2016–17—a 2.7% increase over last year.
Campus Usage
SDS courses serve a broad cross-section of students across campus. Enrollment in undergraduate SDS courses during 2016–17 is shown by college:

Enrollment in Undergraduate SDS Courses by College

Enrollment in graduate SDS courses during 2016–17 is shown by college:

Enrollment in Graduate SDS Courses by College
Statistics Online Support (SOS)
Dr. Kristin Harvey and Sally Ragsdale were awarded a Curriculum Innovation Grant from the Center for Teaching and Learning in Spring 2015 that funded the creation of the Statistics Online Support (SOS) resource platform. Designed for student researchers who have collected data but need guidance selecting and performing statistical analyses, SOS provides on-demand content and tutorials that guide users to independently select the data analysis method appropriate for answering their research questions. Specifically, Dr. Harvey and Ms. Ragsdale created a website that provides users a structured decision tree based platform where users follow embedded links that guide them to the types of statistical methods appropriate for their situation and tutorial videos on how to carry out those analyses.

The site was made publicly available in Fall 2016 (at https://sites.utexas.edu/sos) and has since seen 915 users have 1,705 sessions on the site with 7,675 page views in total. In Summer 2017, Dr. Harvey and Ms. Ragsdale will create additional content for the site to cover more advanced topics such as mediation in regression and logistic regression.

Ms. Ragsdale presented the site and features to instructors for the undergraduates in Freshman Research Initiatives (FRI) and Scientific Inquiry Across the Disciplines (SIAD) signature courses in Fall 2016 and received positive feedback on the usefulness of the site for students. With the addition of more advanced topics, the site can be used for graduate students working on quantitative research related to their dissertations as well as journal article submissions. Participants in several courses at SSI, including UT staff and faculty, were also introduced to the site and were excited to begin using the tool as they conduct their own research.

MOOC:
In September 2016, Dr. Michael J. Mahometa released a third run of the department’s Massively Open Online Course (MOOC), Foundations of Data Analysis (FoDA), through the edX platform. This version is the most stable version to date, providing online learners with a shortened two-part course: one for Descriptive Statistics and Visualizations, and another for Inferential Statistics, each six weeks in length, with live office hours each week hosted through Google+ Hangouts OnAir. FoDA includes 12 total weeks of instruction, with over 40 instructional content videos, over 20 R tutorial videos, over 120 direct feedback videos for incorrect answers, and a scaffolded learning structure that reliably guides students through the analysis process, while increasing independence.

This third offering, like its predecessors, showed remarkable success. Almost 20,000 students, representing 161 countries, enrolled in both parts of FoDA. The course also saw over 30% participant retention in Part One and over 50% participant retention in Part Two, based on video views in the last week of each part.

OnRamps
Since Fall 2013, Dr. Michael Mahometa and Sally Ragsdale have worked alongside Dr. Cathy Stacy, and with the Center for Teaching and Learning, to transform SDS 302 Data Analysis for the Health Sciences into a course for The University of Texas OnRamps initiative. This initiative continues to serve a wide variety of Texas High School students, offering dual enrollment for college credit in SDS 302. Twenty-six Texas High Schools participated in the initiative during the 2016/2017 High School Year, with 666 students enrolled in the Dual Enrollment course. Of those students enrolled, over half (56.16%) were eligible for college credit, with almost 30% (195 students) accepting their UT SDS 302 college credit.
Noteworthy:

- Dr. Mahometa’s edX MOOC (Foundations of Data Analysis, parts 1 and 2) has been approved for continued deployment for the 2017/2018 academic year.
- Analysis of the second FoDA MOOC data showed a significant improvement for students exposed to the direct feedback videos. Students who received direct video feedback on incorrect PreLab questions showed significantly better subsequent performance on Labs, as compared to students who did not received direct video feedback.
B. CONSULTING

SDS provides free statistical consulting services to students, faculty, and staff. Clients receive assistance in planning and interpreting analyses, working with statistical software, developing research study designs, and learning how to better organize and manipulate their data. Faculty may also request contract consulting services for more in-depth analyses. In addition, SDS offers a collection of highly-subscribed instructional short courses in various statistical methods and software packages for a nominal fee.

Free Consulting
Free consulting services continued to be in high demand in 2016–17. SDS consultants provided 987 hours of free consulting to UT faculty, staff, and students, for an average of 19.7 hours per week. Student clients used 81% of these hours, while faculty and staff used the remaining 19%. Ninety-eight percent of consulting clients reported that their most recent consult was either “very good” or “excellent.” Nearly all 2016–17 consulting clients (99%) reported that they felt more confident about doing research and that their research results will be of higher quality because of the consulting service.

![Free Consulting Service by College](chart)

Contract Consulting
SDS provided 146 hours of contract consulting to UT faculty and organizational units and the Texas State Board of Podiatric Medical Examiners in 2016–17. Contract consulting clients generally have more complex data analysis needs and pay a fee for these services. SDS consultants typically provide a detailed reporting of results suitable for publication. Contract consulting clients during the 2016–17 academic year included faculty and staff from the College of Natural Sciences, Department of Neuroscience (32%); Dell Medical School (31%); the College of Education (7%); the School of Law (7%); Moody College of Communication (7%); the Office of Research Support (7%); state agencies (7%); and the College of Liberal Arts, Department of Linguistics (3%).
Short Courses
In 2016–17, SDS taught a total of 27 short courses, covering various software packages such as R, Stata, Python, and Matlab, to 559 registrants. SDS offered two new SQL short courses in Fall 2016 and continued offering the Topics in Statistics series taught by the statistical consulting staff.

Overall, SDS’s software and topic short courses saw high enrollment and positive student feedback in 2016–17. Ninety-four percent of all short course participants said they would recommend the course they took to others.

Noteworthy:
• SDS Consulting services continued to be in high demand across the University, with over 99% of all 2016–17 available appointments filled and an average wait time exceeding 10 days.
• Graduate student Joy Wyckoff joined the consulting team in Spring 2017 as a student consultant. Joy was a Graduate Fellow during Fall 2016.
C. SDS GRADUATE DEGREE PROGRAMS

PhD in Statistics
The PhD in Statistics program welcomed its fourth cohort of five students in August 2016 and opened its application window for Fall 2017 in September 2016. The program received 125 completed applications for Fall 2017 admission. Twenty offers of admission were made, resulting in an expected yield of seven students.

Admissions Data: PhD Statistics

<table>
<thead>
<tr>
<th></th>
<th>Fall 2013</th>
<th>Fall 2014</th>
<th>Fall 2015</th>
<th>Fall 2016</th>
<th>Fall 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applied</td>
<td>79</td>
<td>104</td>
<td>130</td>
<td>138</td>
<td>125</td>
</tr>
<tr>
<td>Admitted</td>
<td>13</td>
<td>13</td>
<td>15</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>Selectivity (%)</td>
<td>16.46</td>
<td>12.5</td>
<td>11.5</td>
<td>10.9</td>
<td>16</td>
</tr>
<tr>
<td>Enrolled</td>
<td>5</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>7*</td>
</tr>
<tr>
<td>Yield (%)</td>
<td>38.46</td>
<td>46.15</td>
<td>53.33</td>
<td>33.33</td>
<td>35.00</td>
</tr>
</tbody>
</table>

* Expected

PhD Student Progress

<table>
<thead>
<tr>
<th>Academic Year</th>
<th>Enrollment Count</th>
<th>Candidacy Count</th>
<th>Graduates</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016–2017*</td>
<td>20</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>2015–2016</td>
<td>16</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>2014–2015</td>
<td>11</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2013–2014</td>
<td>5</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*Enrollment and graduation numbers current through Summer 2016

Master’s of Science in Statistics
The M.S. in Statistics program only accepted applications from current UT graduate students who are applying to earn the degree concurrently with a doctoral degree in another discipline such as Government, Educational Psychology, Civil Engineering, and Sociology.

Enrollment Trends

<table>
<thead>
<tr>
<th>Academic Year</th>
<th>Enrollment Count</th>
<th>Graduates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M.S. Only</td>
<td>M.S./PhD</td>
</tr>
<tr>
<td>2016–2017*</td>
<td>7</td>
<td>31</td>
</tr>
<tr>
<td>2015–2016</td>
<td>10</td>
<td>37</td>
</tr>
<tr>
<td>2014–2015</td>
<td>28</td>
<td>25</td>
</tr>
<tr>
<td>2013-2014</td>
<td>22</td>
<td>18</td>
</tr>
<tr>
<td>2012-2013</td>
<td>30</td>
<td>17</td>
</tr>
<tr>
<td>2011-2012</td>
<td>28</td>
<td>10</td>
</tr>
<tr>
<td>2010-2011</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Enrollment and graduation numbers current through Spring 2017
Noteworthy:

- Dr. Oscar Madrid Padilla is the first PhD graduate of the program. Oscar defended his dissertation in May 2017. He worked on his dissertation, *Constrained Estimation via the Fused Lasso and Some Generalizations* under the supervision of Dr. James G. Scott. In July 2017, he will start a postdoctoral fellowship at the University of California–Berkeley working with Dr. Michael Jordan.
- Five PhD in Statistics students entered candidacy during Fall 2016: Omar Chavez, Maurice Diesendruck, Novin Ghaffari, Bowei Yan, and Michael Zhang.
- Four Professional Development (travel) Awards were provided to current PhD in Statistics students to support a paper or poster presentation of their research. Funding for these awards was provided by the Graduate School and/or the department. The recipients of the awards were:
  - Carlos Pagani Zanini: to present a poster, “A Model for Sequential Refinement and Coagulation of Random Partitions” at the Conference of Texas Statisticians 2017 in Dallas, Texas.
- Two PhD students have publications or accepted publications:
D. GRADUATE FELLOWS PROGRAM

The Graduate Fellows program is a selective semester-long mentorship for UT graduate students that provides considerable training and experience in statistical analysis and consulting on applied problems in a variety of disciplines. Students learn new statistical methods and gain the confidence to teach themselves additional methods in the future. The students also build the interpersonal and presentation skills required of a professional consultant.

The Graduate Fellows take the SDS 388 Consulting Seminar taught by Dr. Michael Mahometa with the data analysis component advised by Dr. Kam Hamidieh. The course focuses on teaching the skills of statistical consulting, data analysis, and statistical methods. Additionally, the Graduate Fellows gain hands-on experience applying statistical methods using data from faculty and researchers across campus.

Fall 2016 Graduate Fellows
Selection for the program is a highly competitive process. For the Fall 2016 program, several graduate students from departments all over campus applied for five positions. To be eligible, students must have 30 hours of graduate study or a master’s degree and have a strong applied statistical background, including use of statistical software. The students selected were

- Spencer Fox (Integrative Biology)
- Hyeseung Koh (Advertising and Public Relations)
- Connor Sheehan (Sociology)
- Chang Sun (Statistics and Data Sciences)
- Joy Wyckoff (Psychology)

Data Analysis Clients in Fall 2016
Data analysis is a fundamental component of the program. Fellows work with faculty members to provide assistance with the statistical or mathematical analysis of their research data. This allows students to gain hands-on experience applying statistical methods to real data while providing faculty and researchers across campus a valuable service. The Graduate Fellows worked with faculty members and Arundo Analytics.

Noteworthy:
- During SDS 388, each student was placed on a team that worked on a single project. This allowed students to dig deeper and get more involved with their projects in contrast to past years when students worked on multiple projects for the semester.
- The feedback from Arundo Analytics client was overwhelmingly positive. Our clients Dr. Laura Colgin and Brian Gereke will be acknowledging the students in the paper they will be submitting.
- Joy Wyckoff, one of the Graduate Fellows, was selected to work as a consultant for the department’s consulting services during Spring 2017.
E. PORTFOLIO AND CERTIFICATE PROGRAMS

SDS provides four unique opportunities for students seeking to develop competencies in statistical modeling or scientific computation. The Portfolio in Applied Statistical Modeling and the Portfolio in Scientific Computation are 12-credit programs available to graduate students interested in strengthening and applying these skills to their research area. The Certificate in Scientific Computation & Data Sciences and the Certificate in Applied Statistical Modeling are similar 18-hour programs available to undergraduate students.

Enrollment Trends

Portfolio in Applied Statistical Modeling
Since its inception in Fall 2009, 211 students have been admitted into the program and 98 students have completed the portfolio requirements.

<table>
<thead>
<tr>
<th>Academic Year</th>
<th>Enrollment Count</th>
<th>Completion Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016–2017*</td>
<td>40</td>
<td>12</td>
</tr>
<tr>
<td>2015–2016</td>
<td>40</td>
<td>17</td>
</tr>
<tr>
<td>2014–2015</td>
<td>39</td>
<td>15</td>
</tr>
<tr>
<td>2013–2014</td>
<td>47</td>
<td>16</td>
</tr>
<tr>
<td>2012–2013</td>
<td>54</td>
<td>15</td>
</tr>
<tr>
<td>2011–2012</td>
<td>43</td>
<td>11</td>
</tr>
<tr>
<td>2010–2011</td>
<td>24</td>
<td>11</td>
</tr>
<tr>
<td>2009–2010</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>98</strong></td>
<td></td>
</tr>
</tbody>
</table>

*Enrollment and graduation numbers current through Spring 2017

Participating students come from 11 colleges and schools across campus.

Student Participation Snapshot

<table>
<thead>
<tr>
<th>College/School</th>
<th># students currently enrolled</th>
<th># students completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cockrell School of Engineering</td>
<td>13</td>
<td>20</td>
</tr>
<tr>
<td>College of Communication</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>College of Education</td>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>College of Liberal Arts</td>
<td>6</td>
<td>23</td>
</tr>
<tr>
<td>College of Natural Sciences</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>Jackson School of Geosciences</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>LBJ School of Public Affairs</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>McCombs School of Business</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>School of Nursing</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>School of Pharmacy</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>School of Social Work</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>40</strong></td>
<td><strong>98</strong></td>
</tr>
</tbody>
</table>
Portfolio in Scientific Computation
Since its inception in Fall 2010, 34 students have been admitted to the program and eight students have completed the portfolio requirements.

<table>
<thead>
<tr>
<th>Academic Year</th>
<th>Enrollment Count</th>
<th>Completion Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016–2017*</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>2015–2016</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>2014–2015</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>2013–2014</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>2012–2013</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>2011–2012</td>
<td>7</td>
<td>-</td>
</tr>
<tr>
<td>2010–2011</td>
<td>7</td>
<td>-</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>8</strong></td>
<td></td>
</tr>
</tbody>
</table>

*Enrollment and graduation numbers current through Spring 2017

Participating students come from three colleges and schools across campus.

Student Participation Snapshot

<table>
<thead>
<tr>
<th>College/School</th>
<th># students currently enrolled</th>
<th># students completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cockrell School of Engineering</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>College of Natural Sciences</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>School of Pharmacy</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>10</strong></td>
<td><strong>8</strong></td>
</tr>
</tbody>
</table>

Certificate in Scientific Computation & Data Sciences
Since its inception in Fall 2009, 259 students have been admitted into the program, and 41 students have completed the program.

<table>
<thead>
<tr>
<th>Academic Year</th>
<th>Enrollment Count</th>
<th>Completion Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016–2017*</td>
<td>84</td>
<td>8</td>
</tr>
<tr>
<td>2015–2016</td>
<td>69</td>
<td>7</td>
</tr>
<tr>
<td>2014–2015</td>
<td>69</td>
<td>4</td>
</tr>
<tr>
<td>2013–2014</td>
<td>79</td>
<td>5</td>
</tr>
<tr>
<td>2012–2013</td>
<td>71</td>
<td>6</td>
</tr>
<tr>
<td>2011–2012</td>
<td>51</td>
<td>6</td>
</tr>
<tr>
<td>2010–2011</td>
<td>27</td>
<td>4</td>
</tr>
<tr>
<td>2009–2010</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td><strong>41</strong></td>
</tr>
</tbody>
</table>

*Enrollment and graduation numbers current through Spring 2017
Participating students come from five colleges and schools across campus.

### Student Participation Snapshot

<table>
<thead>
<tr>
<th>College/School</th>
<th># students currently enrolled</th>
<th># students completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cockrell School of Engineering</td>
<td>18</td>
<td>11</td>
</tr>
<tr>
<td>College of Liberal Arts</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>College of Natural Sciences</td>
<td>49</td>
<td>25</td>
</tr>
<tr>
<td>Jackson School of Geosciences</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>McCombs School of Business</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>84</strong></td>
<td><strong>41</strong></td>
</tr>
</tbody>
</table>

### Certificate in Applied Statistical Modeling

Since its inception in Fall 2013, 134 students have enrolled and 33 students have completed the program.

<table>
<thead>
<tr>
<th>Academic Year</th>
<th>Enrollment Count</th>
<th>Completion Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016–2017*</td>
<td>97</td>
<td>23</td>
</tr>
<tr>
<td>2015–2016</td>
<td>56</td>
<td>8</td>
</tr>
<tr>
<td>2014–2015</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>2013–2014</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>33</strong></td>
<td></td>
</tr>
</tbody>
</table>

*Enrollment and graduation numbers current through Spring 2017

Participating students come from six colleges and schools across campus.

### Student Participation Snapshot

<table>
<thead>
<tr>
<th>College/School</th>
<th># students currently enrolled</th>
<th># students completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cockrell School of Engineering</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>College of Communication</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>College of Education</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>College of Liberal Arts</td>
<td>28</td>
<td>9</td>
</tr>
<tr>
<td>College of Natural Sciences</td>
<td>36</td>
<td>16</td>
</tr>
<tr>
<td>McCombs School of Business</td>
<td>20</td>
<td>7</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>97</strong></td>
<td><strong>33</strong></td>
</tr>
</tbody>
</table>

**Noteworthy:**
- Enrollment in the Certificate in Applied Statistical Modeling saw a significant increase this year. Sixty-seven students enrolled in AY 2016–17; equaling the total enrollment since the start of the program in Fall 2013.
F. SDS SEMINAR SERIES

SDS Seminar Series
Since Fall 2011, the SDS Seminar Series has made a vital contribution to the intellectual, cultural, and scholarly environment at The University of Texas at Austin for students, faculty, and the wider community. The lecture series provides participants with the opportunity to hear from leading scholars and experts who work in different applied areas, including business, biology, medicine, computer modeling, and economics. Each talk is free of charge and open to the public. The Fall 2016 SDS Seminar Series featured eight speakers with an average attendance of 30 people. The Spring 2017 SDS Seminar Series featured 16 speakers with an average attendance of 30. (Please see Appendix C for program details.)
G. SUMMER STATISTICS INSTITUTE

The 10th annual UT Summer Statistics Institute (SSI) was held May 22–25, 2017, in collaboration with Academic Technology Support in the College of Liberal Arts building. SSI provides a unique hands-on opportunity for participants to acquire valuable skills directly from experts in the field. Participants joined the Institute from across the country and the world, coming from as far away as Massachusetts and Vietnam, illustrating the growing popularity of the UT Summer Statistics Institute both nationwide and worldwide. SSI featured 23 twelve-hour courses designed to appeal to a broad range of students, faculty, staff, and the public. New courses this year included “Introduction to Meta-Analysis” (Dr. Tasha Beretvas) and “Survival Analysis” (Dr. Bindu Viswanathan).

The 2017 SSI had a participant enrollment of 577 participants with 23 participating instructors from 15 departments. This year’s attendance breakdown was: 32% UT students, 25% UT faculty and staff, 10% non-UT students, and 34% non-UT other (e.g., private industry, state departments, and non-profits).

The 2017 SSI brought in a gross income of $198,920.00 with a projected total net profit of $90,277.00.

Noteworthy:
- 43% of SSI enrollees (248 of the 577 participants) came from outside The University of Texas at Austin.
- 95% of SSI participants who completed the course evaluations said that they would recommend SSI to others.
- 94% of SSI participants who completed the course evaluations said that the instructor increased their knowledge and competence in the area of study.
H. CORPORATE PARTNERSHIPS

During Fall 2016 and Spring 2017, SDS initiated contacts with various companies including Exxon Mobile, EDF Trading, Total, USAA, PPDI, Deloitte Consulting, Arundo Analytics, Infor, Resignation Media, Civitas Learning, and Sysco. USAA and EDF representatives gave presentations and invited students to apply for summer internships. SDS is continuously working to deepen the departments’ connections with the outside partners.
I. GRANTS

During 2016-17, SDS assisted faculty with the submission of more than twelve individual faculty-driven research proposals to such agencies as:

- National Institutes of Health (NIH)
- National Science Foundation (NSF)

and in collaboration with:

- NorthShore University Health System
- University of Pennsylvania
- Harvard University
- University of Florida
- Brown University
- University of Tennessee Health Science Center
- Pennsylvania State University
- University of Texas Health Science Center

SDS had 11 active grants and contracts this fiscal year:

<table>
<thead>
<tr>
<th>Faculty</th>
<th>Active Grants &amp; Contracts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michael Daniels, Professor</td>
<td>• PI: Bayesian Approaches for Missingness and Causality in Cancer and Behavior Studies, NIH R01, ($1,401,344)</td>
</tr>
<tr>
<td></td>
<td>• PI: Predoctoral Training in Biomedical Big Data Science, NIH, ($1,018,560)</td>
</tr>
<tr>
<td></td>
<td>• Sub-award PI: Magnetic Resonance Imaging and Biomarkers for Muscular Dystrophy, NIH, ($118,548)</td>
</tr>
<tr>
<td></td>
<td>• Sub-award PI: Hospital Responses to Medicare Readmission Penalties, NIH, ($84,337)</td>
</tr>
<tr>
<td></td>
<td>• Sub-award PI: Rural Lifestyle Eating and Activity Program (Rural LEAP), NIH/National Heart, Lung, and Blood Institute, ($45,058)</td>
</tr>
<tr>
<td></td>
<td>• Sub-award PI: Non-Parametric Bayesian Methods for Causal Inference, NIH/National Institute of General Medical Sciences, ($413,114.40)</td>
</tr>
<tr>
<td></td>
<td>• Sub-award PI: Optimizing HIV Treatment Monitoring under Resource Constraints, NIH, ($239,026)</td>
</tr>
<tr>
<td>Peter Müller, Professor</td>
<td>• Co-Investigator: Bayesian Approaches for Missingness and Causality in Cancer and Behavior Studies, NIH R01, ($1,401,344)</td>
</tr>
<tr>
<td></td>
<td>• Sub-award PI: Bayesian Inference for Tumor Heterogeneity with Next-Generation Sequencing Data, NIH R01, ($779,800)</td>
</tr>
</tbody>
</table>
Purnamarita Sarkar, Assistant Professor

- PI: *Inference for network models with covariates: Leveraging local information for statistically*, NSF, ($160,000)

James Scott, Associate Professor

- PI: *CAREER: Bringing Richly Structured Bayesian Models into the Discrete-data Realm via New Data-augmentation Theory and Algorithms*, NSF, ($400,000)
- Co-Investigator: *Bayesian Approaches for Missingness and Causality in Cancer and Behavior Studies*, NIH R01, ($1,401,344)
- Co-Investigator: *Predoctoral Training in Biomedical Big Data Science*, NIH, ($1,018,560)

Sinead Williamson, Assistant Professor

- Co-PI: *BIGDATA:F:DKA: Collaborative Research: Theory and Algorithms for Parallel Probabilistic Inference with Big Data, via Big Model, in Realistic Distributed Computing*, NSF ($300,000)

**Noteworthy:**

- Completed first year of BD2K grant: In this first year, curriculum was developed, and five predoctoral trainees successfully completed their training during the inaugural year of the BD2K training grant. This first cohort completed the required core coursework and attended the NIH BD2K consortium meeting in Bethesda, Maryland in November 2016.
- Stephanie Tomlinson joined SDS as the Grants and Contracts Specialist in February 2017. In addition to providing pre-and post-award support for faculty and students, Stephanie will manage the T32 training grant.
- In March 2017, Stephanie Tomlinson gave a presentation to SDS graduate students to provide an overview of external pre-doctorate funding, including fellowships and research grants.
J. DEVELOPMENT

SDS has partnered with the College of Natural Sciences (CNS) External Relations team to create both short- and long-term plans for building philanthropic and community partnerships with alumni, friends, corporations, and foundations. The third annual SDS newsletter was distributed to over 2000 recipients in October 2016 with a follow-up in May 2017. SDS Development personnel met with industry representatives from companies such as Exxon Mobile, EDF Trading, Total, USAA, PPDI, Deloitte Consulting, Arundo Analytics, Infor, Resignation Media, Civitas Learning, and Sysco to discuss partnering opportunities. SDS is continuing to actively pursue opportunities for support of programs and conferences through NIH, NSF, and the Institute of Mathematical Statistics Meeting sponsorship.
Appendix A: SDS Organizational Chart

Dean
College of Natural Sciences
Linda Hicke, Ph.D.

Chair
Mike Daniels, Sc.D.

SDS Full Faculty
(100% Time)
Purnamrita Sarkar, Ph.D.

Assistant Director for Administration
Vicki L. Keller, MA
Coordinates daily SDS operations and student programs
Supervise departmental staff
Graduate Coordinator
Budgets & EOM
Communications
Reporting
FAR

Consulting Manager
Michael J. Mahometa, Ph.D.
Manages all consulting operations
Oversees FAC 101B & Stat App Server
On-ramps & MOOC development

Lecturer +
Kristin Harvey, Ph.D.
Coordinates SDS 302 courses
Teaches undergraduate courses
Course & Curriculum representative

Lecturer +
Kam Hamidieh, PhD
Manages Graduate Fellows program & Corporate Partnership program
Teaches graduate & undergraduate courses

Administration Manager
Sasha Schellenberg, B.A.
Supervise classified staff
Course scheduling & catalog management
CIS, ALEKS, FASET, Final Exam scheduling, HB 2504 admin
Coordinates SDS Seminar Series & Summer Statistics Institute
Event planning

Grants & Contracts Specialist
Stephanie Tomlinson
Grants: pre- and post-award processing
T-32 Grant Management

Senior Administrative Associate
Rachel Poole
HR processing
Purchasing
Payroll
Accounting
Inventory
Records Maintenance
Billing
DEFINE Vouchers

Lecturers
Lauren Blondeau, PhD
Sarah Collins, PhD
Steven Hernandez, MS
Matt Hersh, PhD
Jerry Manheimer, PhD
Maggie Myers, PhD
Mary Parker, PhD
Paul Robbins, PhD
Lindsey Smith, PhD
Bindu Viswanathan, PhD

SDS Core Faculty
(50% Time)
Michael Daniels, Sc.D. Integrative Biology
Peter Müller, Ph.D. Mathematics
Stephen Walker, Ph.D. Mathematics
Carlos Carvalho, Ph.D. IROM
James Scott, Ph.D. IROM
Sinead Williamson, Ph.D. IROM

SDS Core Faculty
(0% Courtesy Appt)
Paul Damien, Ph.D. IROM
Lauren Meyers, Ph.D. Integrative Biology
Tom Sager, Ph.D. IROM
Tom Shively, Ph.D. IROM
Mingyuan Zhou, Ph.D. IROM

SDS Core Faculty
(50% Time)

Grants & Contracts Specialist
Stephanie Tomlinson
Grants: pre- and post-award processing
T-32 Grant Management

Senior Administrative Associate
Rachel Poole
HR processing
Purchasing
Payroll
Accounting
Inventory
Records Maintenance
Billing
DEFINE Vouchers

Student Assistant
Ramzi Saud
Camry Isaac (ULN)
Appendix B: SDS Core Faculty

Carlos Carvalho, Associate Professor, IROM

Dr. Carvalho’s interest is in the development of methodological aspects of structured probability models for large-scale multivariate problems, with applications ranging from financial time series to high-throughput cancer genomics. His work pays special attention to the development and improvement of associated computational tools for model selection and inference, with current projects in financial econometrics and empirical asset pricing problems.

Dr. Carvalho did his undergraduate studies in Brazil. He received his PhD in Statistics from Duke University in 2006. Before moving to The University of Texas, he was an Assistant Professor at The University of Chicago Booth School of Business.

Honors:
- Donald D. Harrington Faculty Fellow – The University of Texas, Austin, 2009–10
- IBM Corporation Scholar – The University of Chicago, 2008–09
- Leonard J. Savage Award for outstanding doctoral dissertation in Bayesian econometrics and statistics – Honorable Mention, 2006

Michael Daniels, Professor, Integrative Biology

Dr. Daniels’ research program revolves Bayesian methods for biostatistics with special attention to incomplete data, estimation of the dependence structures, and methods for causal inference. His current collaborations include clinical trials in weight management (which motivates development of causal methods for mediation) and muscular dystrophy (which motivates development of complex latent variable methods), and questions involving the impact of recent Medicare legislation on ‘preventable’ hospital outcomes.

Dr. Daniels did his undergraduate studies at Brown University and received his doctoral training in biostatistics at Harvard University in the early 1990’s. Most recently, he was Professor and Chair in the Department of Statistics at the University of Florida. Before that, Dr. Daniels spent five years on the faculty at Iowa State University and two years at Carnegie Mellon University.

Honors:
- Editor of Biometrics, 2015–2017
- The Lagakos Distinguished Alumni Award, Department of Biostatistics, Harvard University, 2014
- Fellow of the American Statistical Association, 2007
- Howard Hughes Medical Institute Predoctoral Fellowship in Biological Sciences, 1992–1995
Peter Müller, Professor, Mathematics

Dr. Müller’s interest is in methods and applications of Bayesian inference. More specifically, he is working on nonparametric Bayesian inference, decision problems, and applications to biomedical research problems. Nonparametric Bayesian inference refers to prior models for infinite dimensional random quantities, typically random probability measures. Decision problems include particular clinical trial design and multiple comparison procedures. Other applications that interest Dr. Müller include inference related to dependence structure, specifically graphical models to formalize inference about dependence for high throughput genomic data. Another large area of application is population pharmacokinetic and pharmacodynamic models, which give rise to many good applications that exploit many of his methodological interests.

Dr. Müller’s undergraduate education is from Universität Wien and Technische Universität Wien, Austria. He received his PhD from Purdue University where he worked under Jim Berger on MCMC for constrained parameter problems. He spent several years at the Institute of Statistics and Decision Sciences (ISDS), Duke University, and at M.D. Anderson Biostatistics.

Honors:
- Fellow of the American Statistical Association
- President of the International Society for Bayesian Analysis, 2010

Purnamrita Sarkar, Assistant Professor

Dr. Sarkar works on large scale statistical machine learning problems with a focus on statistical models, asymptotic theory and scalable inference algorithms for large networks.

Dr. Sarkar graduated from the School of Computer Science at Carnegie Mellon University in 2010. After earning her doctorate she was a postdoctoral scholar at U. C. Berkeley jointly in the Department of Electrical Engineering and Computer Sciences and the Department of Statistics. She received her Bachelor’s degree in Computer Science from the Indian Institute of Technology, Kharagpur in 2004.

Honors:
- Best paper award, 29th International Conference on Data Engineering (ICDE), 2013
- Best paper award, International Conference on Social Networks Analysis and Mining (ASONAM), 2009
James G. Scott, Associate Professor, IROM

Dr. Scott's research focuses on statistical methodology for high-dimensional data sets, with applications in a diverse set of areas spanning the social, physical, and biomedical sciences. Three areas of methodological focus include (1) large-scale multiple testing, anomaly-detection and screening problems, where the rate of false discoveries must be controlled in order to yield viable inferences; (2) inference in sparse models; and (3) the application of data-augmentation theory and algorithms to improve the efficiency of Bayesian inference in large-scale models for discrete data sets. His recent applied work has included collaborations in health care, demography, linguistics, biology, and neuroscience.

Dr. Scott received his PhD is from Duke University, where he studied Bayesian model selection under Jim Berger. Before that he studied at Trinity College, Cambridge for two years. He was an undergraduate from 2000 to 2004 at UT-Austin in the Dean's Scholars and Plan II honors programs.

Honors:
- Regents’ Outstanding Teaching Award, 2014
- NSF CAREER Grant, 2013
- Savage Award, 2010 (awarded by the International Society of Bayesian Statistics for best thesis in Bayesian statistical theory)
- National Science Foundation Graduate Research Fellowship, 2006–2009

Stephen G. Walker, Professor, Mathematics

Dr. Walker’s main research focus is on Bayesian parametric and nonparametric methods. He has worked on applications, methodology, theory, implementation via MCMC, and foundational issues. Dr. Walker’s main areas of applications include medical statistics and financial data. Recent work on Bayesian nonparametrics includes constructing time series and regression models. Recent work also includes working with Bayesian models under misspecification and using loss functions as an alternative to probability models within a learning process akin to Bayesian updating.

Dr. Walker received his BA (Hons.) in Mathematics at the Oriel College of Oxford University, being awarded Open Exhibition on entry to the college. He received his PhD in Statistics from the Imperial College of London in 1995, supervised by Jon Wakefield. Dr. Walker has taught at various institutions: Imperial College at London, the University of Bath, and most recently at the University of Kent.

Honors:
- Editor of Journal of Statistical Planning and Inference, 2017–
- Chair of Bayesian Nonparametric Section of ISBA, 2010–2012
- EPSRC Advances Research Fellow, 2001–2006
Dr. Williamson's main research focus is the development of nonparametric Bayesian methods for machine learning applications. In particular, she is interested in constructing distributions over correlated measures and structures, in order to model correlated data sets or data with spatio-temporal dependence. Examples include models for documents whose topical composition varies through time, and models for temporally evolving social networks. A key research goal is the development of efficient inference algorithms for such models, and she is currently investigating methods that allow us to apply Bayesian nonparametric techniques to large datasets.

Dr. Williamson received her MEng from the University of Oxford, MSc from University College London, and PhD from the University of Cambridge. Before joining the faculty at UT Austin, Sinead was a postdoctoral scholar at Carnegie Mellon University.
# Appendix C: SDS Seminar Series Speakers

<table>
<thead>
<tr>
<th>Date</th>
<th>Speaker Name</th>
<th>Institution</th>
<th>Title of Talk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sep. 2, 2016</td>
<td>John Paisley</td>
<td>Columbia University</td>
<td>“Structured and scalable Bayesian nonparametric dictionary learning”</td>
</tr>
<tr>
<td>Sep. 16, 2016</td>
<td>Ramsés H. Mena</td>
<td>Universidad Nacional Autónoma de México</td>
<td>“On some dependent increment processes derived from Bayesian ideas”</td>
</tr>
<tr>
<td>Oct. 10, 2017</td>
<td>Kevin Kindall</td>
<td>EDF Trading</td>
<td>“What Does a Data Scientist Do in an Energy Trading Company?”</td>
</tr>
<tr>
<td>Oct. 28, 2016</td>
<td>Oscar Madrid Padilla</td>
<td>The University of Texas at Austin</td>
<td>“The DFS fused lasso: nearly optimal linear-time denoising over graphs and trees”</td>
</tr>
<tr>
<td>Nov. 11, 2016</td>
<td>Qixing Huang</td>
<td>The University of Texas at Austin</td>
<td>“Visual Correspondences in the Big Data Era”</td>
</tr>
<tr>
<td>Nov. 18, 2016</td>
<td>Tianjian Zhou</td>
<td>The University of Texas at Austin</td>
<td>“PairClone: A Bayesian Subclone Caller Based on Mutation Pairs”</td>
</tr>
<tr>
<td>Jan. 20, 2017</td>
<td>Amy Willis</td>
<td>Cornell University</td>
<td>“Confidence sets for phylogenetic trees”</td>
</tr>
<tr>
<td>Jan. 27, 2017</td>
<td>Rajarshi Mukherjee</td>
<td>Stanford University</td>
<td>“Sparse Signal Detection with Binary Outcomes”</td>
</tr>
<tr>
<td>Jan. 31, 2017</td>
<td>Alexander Franks</td>
<td>University of Washington</td>
<td>“Bayesian Covariance Estimation with Applications in High-throughput Biology”</td>
</tr>
<tr>
<td>Feb. 14, 2017</td>
<td>Matteo Ruggiero</td>
<td>University of Torino &amp; Collegio Carlo Alberto</td>
<td>“Conjugacy properties of time-evolving Dirichlet and gamma random measures”</td>
</tr>
<tr>
<td>Mar. 10, 2017</td>
<td>Guy Cole</td>
<td>The University of Texas at Austin</td>
<td>“Stochastic Blockmodels with Edge Information”</td>
</tr>
<tr>
<td>Mar. 24, 2017</td>
<td>David Van Dyk</td>
<td>Imperial College London</td>
<td>“Quantifying Discovery in Astro/Particle Physics: Frequentist and Bayesian Perspectives”</td>
</tr>
<tr>
<td>Date</td>
<td>Name</td>
<td>Institution</td>
<td>Title</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------</td>
<td>----------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>Mar. 31, 2016</td>
<td>Xi Chen</td>
<td>NYU Stern</td>
<td>“Statistical Inference for Model Parameters with Stochastic Gradient Descent”</td>
</tr>
<tr>
<td>Apr. 3, 2017</td>
<td>Robert Tibshirani</td>
<td>Stanford University</td>
<td>“Some Progress and Challenges in Biomedical Data Science”</td>
</tr>
<tr>
<td>Apr. 14, 2017</td>
<td>Peter Hoff</td>
<td>Duke University</td>
<td>“Adaptive FAB confidence intervals with constant coverage”</td>
</tr>
<tr>
<td>May 5, 2017</td>
<td>Juhee Lee</td>
<td>UC Santa Cruz</td>
<td>“Joint Bayesian Semiparametric Regression Analysis of Recurrent Adverse Events and Survival”</td>
</tr>
</tbody>
</table>