E-Tidings Newsletter

SCASA Events and News

April, 2026

Volume 14, Issue 4

In This Issue

SCASA's 45th Applied Statistics Workshop * Friday, May 8, 2026

Dr. Enju Liu

MD, PhD

Senior Biostatistician, Institutional Centers for Clinical and Translational Research (ICCTR)

Assistant Professor of Pediatrics, Harvard Medical School

SCASA cordially invites you to the 45th Annual Applied Statistics Workshop

Date: Friday, May 8, 2026

Time: 9:00 AM to 2:15 PM (PST)

Place: on Zoom

Instructor: Dr. Enju Liu

Topic: Understanding Collider Bias in Biostatistical Analysis

Cost: Student (Undergraduate, Graduate, Post-doc): $5 ■ SCASA Member (SCASA/OCLB/SDASA): $10 ■ Non-SCASA Member: $15

Registration: Link

Zoom Link: Will be sent to registrants

Continues on the next page.

Continued from the previous page.

Instructor's Bio: Dr. Enju Liu is an Assistant Professor in Pediatrics at Harvard Medical School. She has a broad background in Medicine, Biostatistics and Epidemiology with specific training in nutrition and statistical methods. She earned her PhD in Nutritional Epidemiology and has extensive experience in design, analysis and conduct of observational studies and randomized controlled trials. She collaborates with investigators at Boston Children's Hospital, across the Harvard Community as well as outside Harvard on the international level. She currently serves as the Associate Director for the Harvard Catalyst Biostatistics Program. Her research has focused on preventive and therapeutic effects of dietary factors, nutritional supplements, and other interventions on child and maternal health outcomes. Her research activities have resulted in over 120 publications in peer-reviewed scientific journals, including the New England Journal of Medicine, Cell, and the American Journal of Clinical Nutrition.

Course Description: In biostatistical analysis, adjustment for potential confounders is a common practice to determine the effect of an exposure or treatment on health outcomes. However, little attention has been given to the potential distortion of the association between exposure and outcome caused by collider bias. A collider refers to a variable that is caused by both the exposure and the outcome or risk factors of the outcome. It is often overlooked but can introduce bias into data analysis, leading to erroneous conclusions. Various statistical approaches commonly employed to control potential confounding, such as restriction, stratification, or adjustment for the collider in regression models, can inadvertently introduce collider bias. This short course aims to equip researchers, biostatisticians, and data analysts in the fields of public health or clinical research with valuable knowledge to increase their awareness of collider bias. Throughout the course, participants will learn how to identify and interpret causal diagrams and directed acyclic graphs (DAGs), which serve as powerful tools for assessing and understanding collider bias. Real-world examples and case studies will be utilized to illustrate the potential impact of collider bias on study results, emphasizing the importance of accurately addressing this bias in data analysis and decision-making.

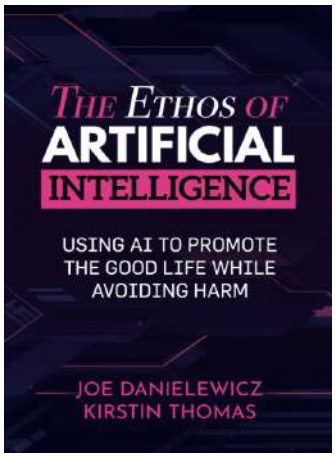
Continues on the next page.

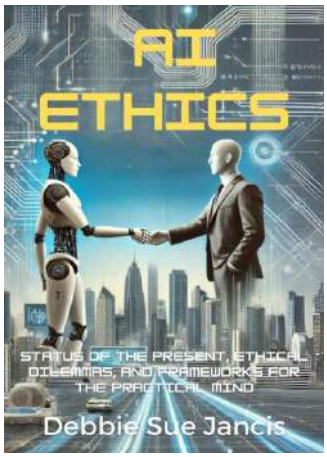
Continued from the previous page.

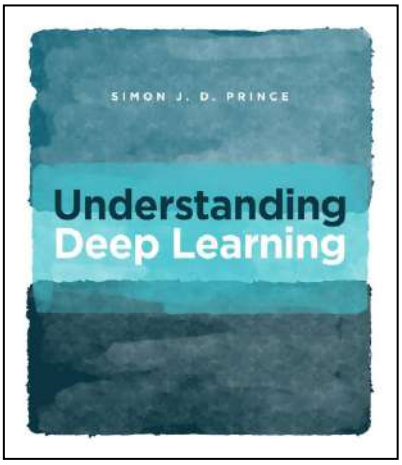
Event Agenda (Pacific Time):

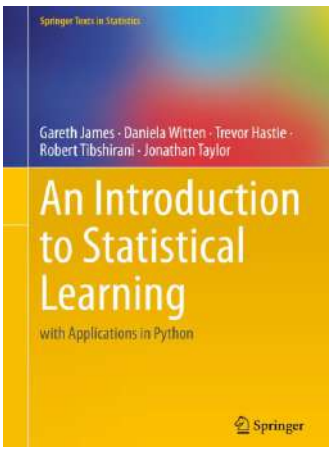
9:00 - 9:10AM Welcome and introduction of speaker
9:10 - 10:00AM Session 1
10:00 - 10:10AM Break
10:10 - 11:00AM Session 2
11:00 - 11:10AM Break
11:10 - 12:00PM Session 3
12:00 - 1:10PM Lunch + **12:20-12:35PM** presentation titled *'Sephora Kids': Reality or Media Hype?* by Tara Annie Kurian (11th grader, University High School, Irvine) – 1st place winner of the 2026 Annual Regional Data Visualization Poster Competition
1:10 - 2:00PM Session 4
2:00 - 2:15PM Q&A and concluding remarks + book raffle (must be present to win)

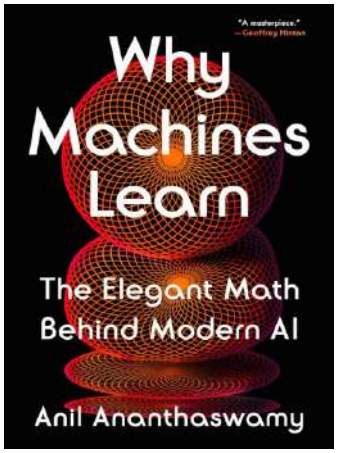
Sneak Peek at Upcoming Raffle Books (x2 each):











REGIONAL DATA VISUALIZATION POSTER CONTEST WINNERS ANNOUNCED!

We are pleased to announce the winners of this year's Poster Contest. After careful review by our judges, the following participants were selected for their outstanding work:

- **First Place:** Tara Annie Kurian (Grade 11, University High School, Irvine) / *'Sephora Kids': Reality or Media Hype?*
- **Second Place:** Om Lala (Grade 11, Whitney High School, Cerritos) / *Bayesian Logistic Regression for Parkinson's Disease Diagnosis*
- **Third Place:** Aireen Kim (Grade 11, Northwood High School, Irvine) / *Is Gen-Z "Brainrot" Real?*
- **Fourth Place:** Rosalinda Ordonez (Grade 9, River City High School, Sacramento) / *Car Accident*
- **Fifth Place:** Isabella De Melo Grade 9, River City High School, Sacramento) / *Model of ADHD*

Congratulations to all our winners, and thank you to everyone who participated and contributed their creativity and effort. The five posters have been submitted to the national-level contest. We will keep our fingers crossed.

The posters can be viewed here:

HTTPS://COMMUNITY.AMSTAT.ORG/SCASA/POSTERCOMP

Regression Modeling Strategies

Professional Development Opportunity

We're excited to share **two** professional development opportunities available this May through the ASA/Instats partnership.

Regression Modeling Strategies - 4-day course

May 14, 15, 18, 19, 2026, 9am–4pm CDT = 7am–2pm PDT

This course has been a cornerstone of statistical education and for the first time, it is available through ASA. Taught by Frank Harrell, PhD (Vanderbilt University School of Medicine) and Drew Levy, PhD (GoodScience, Inc.), it offers rigorous statistical methodology for reproducible research. Emphasis is on developing predictive models, ordinal semiparametric models, model interpretation and validation, and quantifying predictive accuracy, plus many more topics including data reduction (unsupervised learning), navigating the choice of statistical models vs. machine learning, causal thinking in model specification, and an introduction to Bayesian modeling. Prerequisite: Some familiarity with multiple linear regression.

Introduction to R, RStudio, Regression, and the rms Package (Pre-RMS) - 1-day course

May 11, 2026, 9am–4:30pm CDT = 7am–2:30PM PDT

For an introduction to R/RStudio & an efficient comprehensive R workflow taking advantage of 35 years of R/S experience, to learn about the R rms package, or to enhance your multiple regression skills to be ready for the Regression Modeling Strategies 4-day course. Taught by Frank Harrell, PhD, this course is designed to get you up to speed with R/RStudio, an efficient modern R workflow, and the rms package. Drew Levy, PhD serves as moderator.

Bundle and save: Register for both courses together and receive a 10% discount.

To learn more about the courses, visit **hbiostat.org/course.**

The Traveling Course is coming to SCASA

Tentative Day/Time: Saturday, October 24, 2026, 9AM-1PM (PST)

Presentation mode: Online

Course Title: Tree-Based Machine Learning Methods for Prediction and Variable Selection

Instructors: Drs. Hemant Ishwaran and Min Lu, University of Miami

Abstract: Tree-based machine learning methods offer several benefits in data analysis, including non-linearity, robustness, scalability and handling mixed data types. This course emphasizes practical learning with hands-on code examples and result interpretations, which is essential for understanding and applying these techniques. Based on the widely popular R package "randomForesSRC", we will present methods for computing predicted outcomes, variable importance indices and other inference estimates. In addition, we will introduce a new model-independent variable selection method, called the rule-based variable priority, and present its implementation using the R package "varPro". For all these analyses, we will cover different types of outcomes including continuous, categorical, multivariate, survival and competing risk outcomes. Utilizing real-world datasets from medicine and public health, topics in these analyses will provide hands-on code, working examples and result interpretations. We will provide additional code for visualizing model results and constructing coefficient tables for interpretation, and address scenarios such as imbalanced classes, unsupervised problems, fast implementation on big data and protection of confidential data.

Special Television Episode Released: *Statistical 'Therapy' – A Thought Experiment & Television Game Show*

We are pleased to announce the release of "Statistical 'Therapy' - A Thought Experiment & Television Game Show," a special feature that explored innovative applications of 'Post-Conventional' Psychology, an interactive program broadcasted on Pasadena Media's Arroyo Channel.

This episode examined hypothesized and observed results of applying statistical tools and philosophies to human moods in mental health, critically focusing on three points: 1.) Review of an interactive game show in which experts and students applied statistical reasoning to mental health case examples. 2.) Overview of conceptual applications of statistical philosophies and tools (outside the scope of formal psychotherapy) for mental health 3.) A concluding segment presenting this work as part of preliminary pilot research, analysis, and open-source review.

Broadcast & Streaming Information

The full episode aired on Pasadena Media's Arroyo TV Channel in March and now the full episode can be found here: **https://youtu.be/a3r50cOKRnE?si=zWiilRxN61pJQmVF**.
Audience engagement and public discussion are warmly encouraged.

Game Show Contestants

· Dr. Yuan Chai (PhD, MA), Linguistics & Phonetics Academic
· Dr. Neil Patel (MD, MBA), Healthcare Business Analyst
· Ms. Mariela Gallo (MPH, MBA, CHES), Hospital Program Manager & Educator
· Ms. Reighan Grandolfo (BS), Applied Biostatistics Scholar, Claremont Graduate University

Show Judges

· Dr. Ronald (Ron) Wasserstein, Executive Director, American Statistical Association
· Dr. Chong Ho (Alex) Yu, 2025–2026 President, Southern California Chapter of the ASA
· Mr. Edward Pineda, Mathematics Professor, Los Angeles City College
· Ms. Veronica Parent, Statistics Student, Los Angeles City College

Production Team

· Dr. Anna Yu Lee (2023–2024 President, SCASA)
· Jasiri Jenkins-Glenn, Sharon Huang, Mark Saikaly, Cliff Present, Ian Menzies (Pasadena Media / The Arroyo

New eBook Release: *R for SAS Programmers: SQL Processing*

by Sunil Gupta

We're excited to announce the release of a new eBook designed specifically for SAS programmers looking to expand their skills into R:

R for SAS Programmers: SQL Processing

This is the first installment in a new series focused on bridging SAS and R for SQL workflows. Built from the foundation of a proven R-Guru mentoring and training program, this book transforms hands-on teaching experience into a structured, accessible learning journey.

If you've been curious about R but unsure where to begin, this eBook offers a clear, guided path. You'll move step by step through familiar SQL concepts, translated into the R environment—making the transition smooth, practical, and efficient.

Why this book stands out:

- Gain a clear understanding of the **key similarities and differences** between SAS SQL and R SQL syntax
- Learn through **visualized tasks and functions** that accelerate comprehension and application
- **Save time coding** with well-organized summaries and practical, ready-to-use examples
- Practice with **100+ task-based exercises** that reinforce real-world skills
- Strengthen your ability to **troubleshoot common programming and data challenges**

Whether you're a seasoned SAS programmer or just starting to explore R, this eBook provides a focused and approachable way to get your feet wet—without unnecessary complexity or frustration.

Start your journey into R with confidence and clarity. Click below.

https://www.lulu.com/shop/sunil-gupta/r-for-sas-programmers-sql-processing/ebook/product-m2g2pr8.html?q=r%2Bfor%2Bsas%2Bprogrammers&page=1&pageSize=4

Dr. Normalcurvesaurus, Ph.D., Presents