

E-Tidings Newsletter



SCASA Events and News

Volume 11, Issue 3

March, 2023



PAGE 2

SCASA Job Fair Report

PAGE 3

SCASA Regional Poster Competition Winners

PAGES 4-6

Applied Statistics Workshop

PAGE 7

Planning ISEF 2024

PAGE 8

Book Review

PAGE 9

Dr. Normalcurvesaurus, Ph.D.



SCASA JOB FAIR REPORT

By Eric Kawaguchi, USC, SCASA VP for Professional Affairs

We had one employer present (LA Dodgers) and two students attended (8 were registered according to EventBrite). Both students were from Cal Poly SLO as Math/Stat majors, one was a senior about to graduate and the other was a junior looking for an internship.

I started reaching out to companies and potential employers two months prior to the job fair. I asked Neal for his contacts. He only provided one, as the others had moved on from their previous employers (something that is not unusual for tech companies). The one I reached out to never responded. I also asked Linda for any possible leads and she provided me two, neither of which have returned emails as well.

Personally, I reached out to City of Hope and Kaiser Permanente, but they only had advanced level positions (MS or PhD required). Most biostat employers (a majority of my connections) require an MS/PhD and given that the job fair was tailored to undergraduates, I decided not to reach out to them. I sent out a few "Call for Employers" on my LinkedIn page but only had one response, a colleague of mine from UCLA that now works for the Dodgers.

Moving forward here are some ideas that I have to improve the job fair:

1. Needing more individuals in industry to help with recruitment of employers. As someone in academia and in biostatistics, most of my connections are limited. It would be helpful to have an industry statistician or two help with recruiting employers as they potentially have a larger network than myself. I will be going to DataFest as a volunteer and will try to connect/network with some of the vendors there.
2. Mock Job Interviews/Resume critiques. In addition to actual recruitment I like the idea of holding mock job interviews / resume critiques. Again, I do not think that I am well equipped for this since my experience is limited as I'm in an academic setting and don't personally hire data analysts/data scientists. I know the ASA holds workshops for graduating students during JSM. I might attend one this year and ask questions to see how we can implement one on our end.
3. Surveys to reach out to students. In general, I think it would be a good idea to see who views the newsletters and attends SCASA events. I have been going to SCASA events since 2013, when I was an undergraduate at CPP; but I know that there aren't many students who do. We probably need to engage the student body more. I personally blast SCASA events to USC biostat students via Slack. The surveys can also ask questions related to the job fair so we know who is interested and for what type of position (and can help tailor the job fair).

I will be happy to have a Zoom chat to de-brief more in person. I can be reached at

eskawaguchi@gmail.com

3RD SCASA REGIONAL STATISTICS DATA VISUALIZATION POSTER COMPETITION WINNERS ANNOUNCED

In March, SCASA hosted its **3rd Regional Statistics Data Visualization Poster Competition** for grades 10-12. The top five winning teams are as listed below. The winning teams will receive cash prizes, medals, and certificates and the top five posters will be submitted to the national-level Data Visualization Poster Competition organized by the American Statistical Association. Wish them good luck!



First place: (\$200) "Covid on the Border" by Raymond Levine, *Westview High School, San Diego*



Second place (\$150) "Sports-Related Mental Imagery Utilization Among High School Student-Athletes" by Lauren Chai, *Diamond Bar High School, Diamond Bar*



Third place (\$100) "A Novel Super-Resolution AI Engine and Multi-Stage Abnormality Detection Pipeline for Early Blindness Prevention" by Justin Liu, *Palos Verdes Peninsula High School, Rolling Hills Estates*



Fourth place (\$50) " Designing a Customizable Fidget Toy that Lowers Motor Tic Frequency" by Mandy Bautista, Kiana Bingcang, Megan Peters, Abraham Rock, *Sato Academy of Mathematics and Science, Long Beach*



Fifth place (\$50) " Positive effect of verbal and visual cues during training enhances sprinting efficiency of the Sato basketball team" by Athan Sarantakis, Tristan Luedee, Christopher Raniag, Rayaan Sakhawat, *Sato Academy of Mathematics and Science, Long Beach*

The 42nd SCASA Annual Applied Statistics Workshop

FRIDAY, APRIL 21, 2023, 9 AM – 4:15 PM

via ZOOM



CLOUD COMPUTING AND SAGEMAKER

When: Friday, April 21, 2023, 9 AM – 4:15 PM

Location: via Zoom (<https://csulb.zoom.us/j/87964831837>)

Cost: Free

Registration: SageMaker account is not guaranteed if not pre-registered at <https://www.eventbrite.com/e/cloud-computing-and-amazon-sagemaker-tickets-585968426337>

About the Presenters:



Gauthier Kelly is a Solutions Architect at Amazon Web Services (AWS). He is a certified expert of multiple AWS cloud computing products, including AWS Partner: Cloud Economics, AWS Certified DevOps Engineer, AWS Certified Solutions Architect.



David Laredo Razo is R&D Solutions Architect at Amazon Web Services (AWS) where he has helped develop multiple machine learning prototypes. His R&D projects include intelligent HVAC, model selection in machine learning, value actuator, natural language processing, time series, and end-to-end ML.



Josh Dunlap is an AWS Account Manager. Josh has spent the last 15 years working alongside Higher Education Universities and currently supports a number of colleges in Southern California.

ABSTRACT

This workshop will introduce participants to the basic concepts and procedures of cloud computing.

Cloud computing is a technology that has revolutionized the way people use computer resources. Cloud computing enables users to access a wide range of software, hardware, and storage resources over the internet, without the need for on-premises infrastructure. Cloud computing skills are in high demand in the job market. As more companies move their infrastructure to the cloud, people who have cloud computing skills will have a competitive advantage when looking for employment. Currently, the market leaders of cloud computing are Amazon Web Services (AWS), Microsoft Azure (partnership with SAS Institute and Open AI), and Google Cloud. In addition to AWS, Amazon SageMaker, which is a cloud-based machine learning tool for predictive modeling, will also be introduced. SageMaker enables data scientists to easily create, train, and deploy machine learning models into production-ready hosted environments. No prior knowledge or experience of cloud computing is required to attend this workshop.

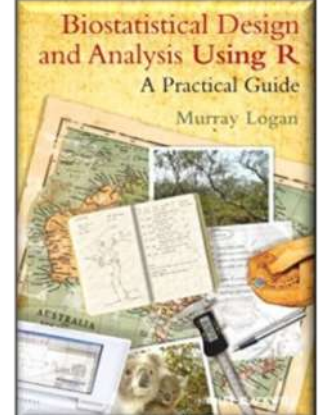
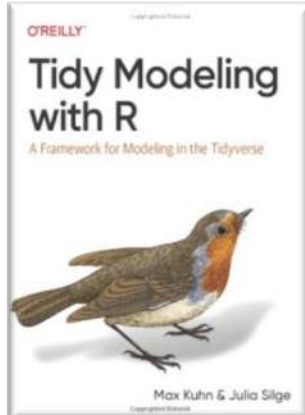
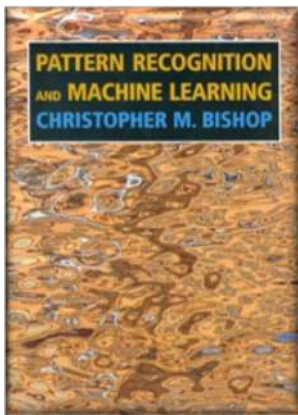
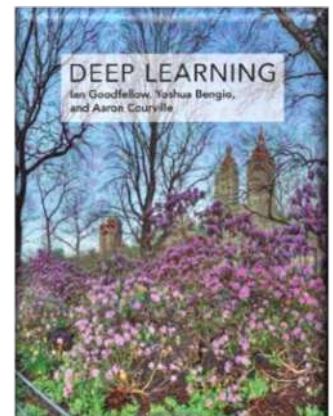
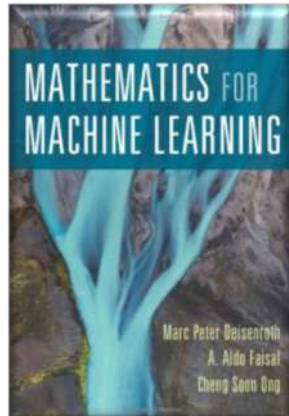
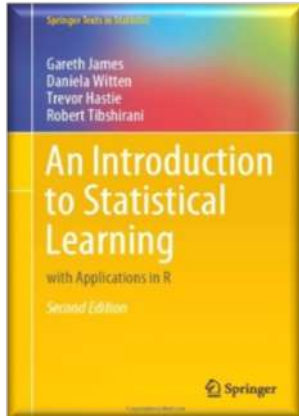
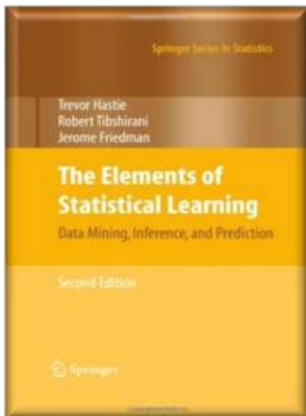
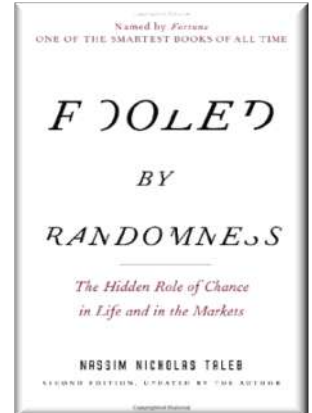
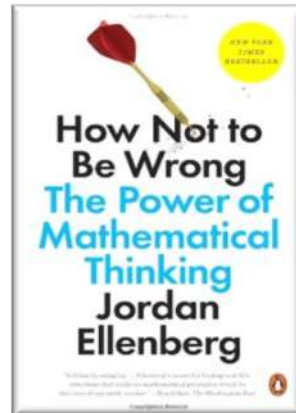
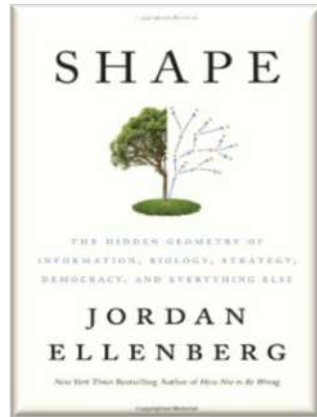
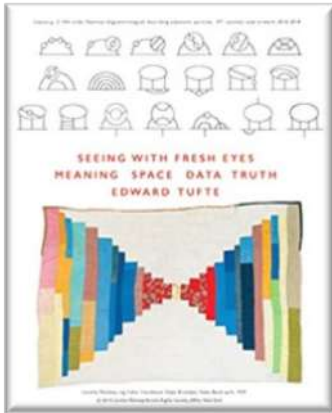
EVENT SCHEDULE

- 9:00 AM – 9:45 AM** Introduction and Overview
- 9:45 AM – 12:00 PM** Session I: Machine Learning Basics (No coding)
- 12:00 PM – 1:00 PM** Lunch Break + Presentation by the winning team at SCASA 2023 Regional Statistics Data Visualization Poster Competition (exact time: TBD)
- 1:00 PM – 4:00 PM** Session II: Machine Learning for Data Scientists
- 4:00 PM – 4:15 PM** Book Raffle and Wrap-up





BOOKS FOR RAFFLE



Regeneron ISEF 2024 ASA Special Awards in Statistics May 11-17 in **Los Angeles!**

By Madeline Bauer, USC (retired)

- **In-person** will (most likely) take place on **Tuesday May 14** and **Wednesday May 15**
- **Virtual judging (IF exists)** would (most likely) be held the weekend of **May 3-6**.

You are invited to participate in the statistics judging and welcome to help with planning for the ASA Special Awards in Statistics at ISEF 2024:

[Click here to indicate your interest:](https://bit.ly/ISEF2024SaveTheDate)
<https://bit.ly/ISEF2024SaveTheDate>

**HELP! "IN SEARCH OF" A STATISTICS TEXTBOOK
THAT INCLUDES ANCOVA FOR EVALUATION OF "PRE-TEST/POST-TEST" DATA!**

By Madeline Bauer, USC (retired)

Many high school science fair projects (properly) use a pre-test/post-test design but then (oh no!) they use a t-test on mean percent change. Mentoring these students, 1-on-1 virtually, I describe the more efficient ANCOVA (<https://hbiostat.org/bbr/change.html>, <https://twitter.com/stephensenn/status/1224362916423573504>).

There is an excellent book by Stephen Senn (<https://www.wiley.com/en-gb/Statistical+Issues+in+Drug+Development%2C+3rd+Edition-p-9781119238577>) that covers this topic but I'm looking for an introductory statistics textbook that covers this useful topic for high school teachers.

Any suggestions are appreciated. Please use this google link form:

<https://forms.gle/MzQbYpNBV5GXjrAi7>



BOOK REVIEW

By Sunil Gupta, Founder of R-Guru.com and SASSavvy.com



I have the honor to review "Data Wrangling with R" by Gustavo Santos, Packt Publishing, February, 2023 ! Being new to R, I was afraid this book was going to be too technical and only on statistical modeling. On the contrary, I was happy to find this R book easy to read, understand and apply whether you are new or an intermediate R programmer. The logical flow between chapters and sections with useful summaries and tips showcase Gustavo's in-depth knowledge and expertise in R. I really like how Gustavo compares similar R methods which help to reinforce R syntax understanding. Gustavo's approach will be liked by Data Scientists since all traditional methods of data cleaning, structure and operations before statistical modeling are followed. The graph gallery makes it easy to look up the correct R graph code without writing R code from scratch. I highly recommend Data Wrangling with R book to establish a strong foundation of R learning!



About the Author

Gustavo R Santos is a Data Scientist since 2020, working daily with high-end technologies such as R, Python, Databricks, as well as SQL.

Gustavo worked for more than 13 years in the Technology Industry and now he works at a Retail company.

He also maintains a blog in Medium where he shares a lot of his experience and codes with the public, helping people to learn and understand Data Science concepts in an easy way.

Gustavo's academic background is originally in Marketing, but he has a certified Degree from Data Science Academy Brasil, an extension in MIT and currently pursued an specialist MBA title from University of São Paulo.

Product details

Publisher : Packt Publishing (February 23, 2023)
Language : English
Paperback : 384 pages
ISBN-10 : 1803235403
ISBN-13 : 978-1803235400
Item Weight : 1.45 pounds
Dimensions : 7.5 x 0.91 x 9.25 inches

<https://www.amazon.com/Data-Wrangling-transform-visualize-tidyverse/dp/1803235403>

Customer Reviews: ★★★★★ 3 ratings

Paperback
\$44.99 prime

Dr. Normalcurvesaurus, Ph.D. presents

