SCASA Leadership Legacy Corner

Dr. Wayne Smith, CSU Northridge

I earned my Bachelor's degree from California State University (Business Administration) and my Ph.D. from Claremont Graduate University (Information Systems and Technology). I'm on the faculty in the Department of Management in the David Nazarian College of Business and Economics.

Academically, I've helped many faculty at CSU Northridge with their data and statistics-related technology. For many years, I was one of two consultants for the campus faculty to help with software such as XLISP, SAS, SPSS, BMDP, S/S+ and related libraries such as NETLIB. I often helped with data wrangling (e.g., "wide to long", unstructured data, missing data, larger-than-memory computation, dummy/one-hot encoding, etc.) but I also helped faculty with item-response theory (to help analyze student exam questions), multiple and logistic regression, and various convex optimizations.

My first experience at a SCASA event was hearing Stanford Prof. Brad Efron discuss bootstrapping and resampling. The talk/workshop was at USC in about the late 1980's. I was invited by SCASA member Prof. Elizabeth Trybus of California State University, Northridge.

I attended many SCASA events through the years, especially ones involving young people. I have been a judge for the annual High School Poster competition and a mock interviewer for data science and analytics for entry-level college students. I attended many Annual statistics lecturers. Probably my biggest work to date was to contribute as a judge at the International Science and Engineering Fair (High School students) held in Los Angeles every three years (I started in 2014, as I recall). This effort was led by an incredibly talented Madeline Bauer quite ably assisted by Luke Thelen and Harold Dick. I worked with all three of them to help improve the skills, knowledge, and abilities of the judges, the reliability and validity the judging, and I also arranged for Wiley to provide several thousand dollars of books each year to be given to the top student-competitors. I am currently the ASA Liasion for SCASA.

In addition to teaching variety of Accounting, Computer Science, Information Systems, Management, and Analytics courses at CSU Northridge, I had a thriving consulting practice more than three decades, mostly in the Los Angeles area.

Continues on the next page.

Continued from the previous page.

While I also did network engineering, cybersecurity, training, and strategy consulting, the following are some of the more notable engagements related to data and statistics.

My first paid consulting work was in 1978 was doing quarterly descriptive statistics and sales variance analysis for a small Caviar wholesaler. All the work was done on a Tandy Radio Shack Model 1 in BASIC.

Subsequently, I co-taught the Los Angeles Unified School District "in-services" (teacher training) in the summers of 1979 and 1980. I taught teachers from around the District how to use a spreadsheet for grading (Visicalc on an Apple II), and program (BASIC on a PDP-11/70 and Fortran on a Univac 1110). The BASIC programming was useful for not only instructional games but also dynamic quizzes which—using the gradebook data—could be individualized for each student in, say, a Math course. The FORTAN code helped with exam question analysis (point-biserial correlation as I recall).

One of my clients in the mid-1980s was Marquardt, an aerospace company in Van Nuys. I worked in the stress engineering department on the National Aerospace Plane project. This was a classified NASA/DoD project at the time. I used a mainframe computer and then a PC to run stress simulations in hand-coded FORTRAN (for Finite Element Analysis) and a language called SIMSCRIPT (for end-to-end discrete workflows), especially to test core metals in and around scramjet engines to fly at hypersonic speeds (e.g., flying from LA to New York in 2.5 hours or from New York to Paris in, say, 4 hours). We used non-parametric statistics to help evaluate the simulation results.

Another one of my clients in the early 1990's was R.R. Donnelly and Sons. My main task was to supervise a small programming team writing an custom "imposition" (printing details) system to fully print and bind magazines such as Ladies Home Journal and numerous Time Magazing monthlies. I also helped develop a couple of "minimization" optimization batch processes to reduce the downtime for the world's largest printing presses. This work was done in C with C-Plex (now owned by IBM).

Most of the work in the past twenty years is to help faculty (and some students!) with big data faculty issues using Oracle, Postgres, MySQL/MariaSQL, DuckDB, and SQLite. My more recent statistical computing work and help is in R, Python, Julia, and various visualization tools such as Observable. For several years, I was on the steering committee for the original LA R Meetup Group.

In addition to teaching and pro-bono data science work, I now help coordinate and support the High-Performance Computing and related storage/network resources for the entire campus (e.g., the NRP/Nautilus/CENIC AIR systems, the Cloudbank/UC Berkeley CAL-ICOR JupyterHub, and the ACCESS-CI/Jetstream2, et al. system). Much of this work involves Science topics such as Molecular Dynamics and Engineering topics such as Mechanical Engineering. I am active in a few NSF grants and travel occasionally to NSF-sponsored HPC workshops.