

Gaston Sanchez, PhD

gaston.stat@gmail.com
gastonsanchez.com
github.com/gastonstat
linkedin.com/in/sanchezgaston

TEACHING AND ADVISING EXPERIENCE

Lecturer, University of California Berkeley, CA. (Jun 2015 – present)

- Stat 2: Introduction to Statistics.
- Stat 20: Introduction to Probability and Statistics.
- Stat 33A: Introduction to Programming in R.
- Stat 33B: Introduction to Advanced Programming in R.
- Stat 131A: Introduction to Probability and Statistics for Life Scientists.
- Stat 133: Concepts in Computing with Data.
- Stat 151A: Linear Modeling
- Stat 154: Modern Statistical Prediction and Machine Learning.
- Stat 159: Reproducible and Collaborative Statistical Data Science.
- Stat 243: Introduction to Statistical Computing.
- Stat 259: Reproducible and Collaborative Statistical Data Science.
- Stat 375: Professional Preparation – Teaching of Probability and Statistics.

Special Consultant, San Francisco State University, CA. (Jul 2020 – Apr 2021)

- Authored learning materials for the Graduate Opportunities to Learn Data science (GOLD) program.
- Created content for BIOL 806/CHEM 806: Exploratory Data Science for Scientists.

Oniris – National College of Veterinary, Food Science and Engineering, Nantes, France. (Apr 2013 – Mar 2014)

- Lectured Master's course on Latent Variable Path Models.
- Taught workshop on Creating R packages.
- Taught workshops on statistical methods using R.
- Mentored PhD and masters students on statistical programming.
- Reviewed dissertations of PhD students in the Department of Sensometrics and Chemometrics.

Development of curricula for teaching Data Science with R, (2010 – present)

(Available in personal website www.gastonsanchez.com)

- Manipulating Data Tables in R.
- Getting Data from the Web with R.
- Creating R Packages in Windows.
- Handling and Processing Strings in R.

Nielsen Group, University of California Berkeley, CA. (Summer 2012)

- Seminar on Statistical Methods for Genetic Association Analysis.

Program on Reproductive Health and the Environment, UC San Francisco (Fall 2011)

- Workshops on R for data manipulation, graphics, and basic statistical analysis.

Department of Statistics and Operations Research. Universitat Politècnica de Catalunya. (2006 – 2008)

- Tutoring R and statistical programming for graduate students.

RESEARCH EXPERIENCE

Insight Data Science, Palo Alto, CA. (Jun 2014 – Sep 2014)

Data Science Fellow

- Created Web app “www.skincareprods.info” to score personal care products based on health risk.
- Scraped data from the Good Guide and EWG using R, stored database in MySQL.

- Applied Regression Trees with cross-validation for computing health scores of personal care products.
- Front end developed with Python (Flask). Visualization with D3js.

Oniris – National College of Veterinary, Food Science and Engineering, Nantes, France. (Mar 2013 – Mar 2014)
Postdoctoral Researcher

- Built a prototype toolbox in R for multiblock data analysis methods.
- Lead seminars, training workshops, and tutoring on multivariate methods and data analysis with R.
- Created R packages *tester*, *turner*, and *matrixkit* for simplifying development of other packages.

Nielsen Group. University of California, Berkeley. (Jun 2011 – Mar 2013)
Postdoctoral Researcher

- Researched statistical tests used in genetic association studies, and created the R package *Assotester*.
- Developed the R packages *colortools*, *arcdiagram* (arc diagrams) and *DiscriMiner* (discriminant analysis).
- Continued as guest researcher between March 2014 and May 2014.

Biology Scholars Program (BSP). University of California, Berkeley. (Feb 2011 – Nov 2012)
Research Analyst II

- Established the roadmap for studying the scope of the Biology Scholars Program between 1994 – 2010.
- Carried out cluster analysis for segmenting BSP members to target specific audiences and needs.
- Conducted analysis to measure the impact of the program on its members. Applied statistical tests in R to compare the difference of the program members against students outside the program.

Program for Reproductive Health & the Environment (PRHE), UC San Francisco. (Aug 2011 – May 2012)
Research Analyst II

- Collaborated with PRHE's scientists on bibliometric studies on Bisphenol A (BPA) datasets.
- Analyzed data from the biomonitoring study "Chemicals in Our Bodies". Analysis in R with statistical tests to identify threshold levels for different chemicals (metals and PBDA's).

Laboratory of Information Analysis and Modeling, UPC, Spain. (Jul 2005 – Jul 2009)
Statistical Programmer

- Research on advanced methods for structural equation modeling and segmentation analysis.
- Created R package "plsrm" for partial least squares analysis (most popular PLS open source package).
- Carried out a customer satisfaction analysis for clients of Sabadell Bank. Analysis performed in R using structural equation models. Provided recommendations to increase Perceived Brand Image and Customer Loyalty.
- Built path models for the Barcelona School of Informatics Alumni Satisfaction Survey to evaluate perceptions between educational experience and job experience. Findings used to modify the syllabus of Computer Science degrees.

Department of Operations Research, IESE Business School, Barcelona, Spain. (Feb 2007 – Nov 2007)
Research Assistant

- Analyzed data from the International Manufacturing Strategy Survey (IMSS) comprising three separate surveys (1997, 2001, 2005) from more than 1700 companies around the world.
- Drafting analysis with tables and graphs for Prof. Adrian Done's book "Global Trends".

EDUCATION

Universitat Politècnica de Catalunya (UPC), Spain.

PhD in Statistics – Department of Statistics and Operations Research, July 2009.

Dissertation: *PATHMOX Approach: Segmentation Trees in Partial Least Squares Path Modeling*

Universitat Politècnica de Catalunya (UPC), Spain.

DEA in Statistics – Department of Statistics and Operations Research, June 2006.

Project: *Partial Least Squares Path Modeling applied to Segmentation Problems*

Instituto Tecnológico Autónomo de México (ITAM), Mexico.
BS in Actuarial Science - Department of Actuary, June 2002.
Thesis: *Cluster Analysis with Web Mining Applications*.

CERTIFICATES

Universitat Oberta de Catalunya (UOC), Spain.
Certificate in Decision Support Systems, December 2007.

Instituto Tecnológico Autónomo de México (ITAM), Mexico.
Certificate in Brokerage Finance, August 2003.

Instituto Tecnológico Autónomo de México (ITAM), Mexico.
Certificate in Credit and Banking, July 2003.

OTHER PROFESSIONAL EXPERIENCE

Swimming Instructor - *Canyon Pool, El Sobrante, CA*. (Spring 2011)
- Instructor for kids and adults basic swimming and water safety.

Rope Techniques Instructor - *Mexican Alpine Club (CAM), Mexico*. (Jan 2011)
- Taught courses on theory and practice of rope techniques levels I and II.
- Rope Essentials, Descending, Ascending, Hauling, Highlines, Self-rescue.

Outdoors Skills and Rock Climbing - *Mexican Alpine Club (CAM), Mexico*. (Jan - Apr 2010)
- Head instructor of Rock Climbing.
- Responsible for the Rock Climbing sessions of the Mountaineering Course LX.

Climbing and Mountaineering - *Mexican Alpine Club (CAM), Mexico*. (Sep 2000 - Aug 2004)
- Instructor of Outdoor Activities, Rock Climbing and Mountaineering.
- Oversaw 8 semesters of courses at CAM

Junior Analyst - *National Banking and Securities Commission (CNBV), Mexico*. (May 2002 - Aug 2004)
- Supervised and audited credit unions in Mexico.
- Visual Basic (Excel) programming for data collection utilities of projects.
- Management of the Credit Unions Supervision department's datasets: data entry, data dictionary, combining data.