Brian Kim

College Park, MD kimbrian@umd.edu kimbrianj

Education

PhD **University of California, Los Angeles**, Statistics — Los Angeles, CA, Sept 2012 – present

- Dissertation: Population Size Estimation using Multiple Respondent-Driven Samples

Amherst College, Mathematics & Philosophy — Sept 2008 – June 2012

- Graduated with Honors in Mathematics
- Thesis: Using Mode Identification Clustering Methods with Pitch F/X Data

Research Interests

- Data Science Education; Social Network Analysis; Network Sampling Methods; Respondent-Driven Sampling; Population Size Estimation: Capture-Recapture and Multiple List methods; Social Science Applications of Machine Learning

Experience

University of Maryland, Co-Director, Social Data Science Major — College Park, MD, Oct 2021 – present, 4 years 6 months

- Led the development of a new major in Social Data Science (officially launched Fall 2022), including working across departments to build an appropriate curriculum, managing the budget, and setting up infrastructure.
- Developed new courses to be part of Social Data Science, including SDSB233: Data Science for Social Sciences and SDSB326: Python Programming for Social Science.
- Launched a new undergraduate Data Curation Fellowship experience in which students identify, obtain, clean, and curate datasets to be used within social science courses, with a public-facing data resource webpage (https://bsos-data.umd.edu).
- Coordinated across colleges with the College of Information Co-Director to set up advising, teaching workload, and major requirements for students.

University of Maryland, Director of Graduate Studies, Joint Program in Survey Methodology — College Park, MD, Sept 2024 – present, 1 year 7 months

- Oversaw updates to the Master's in Survey and Data Science program, including adding SURV613: Machine Learning for Social Science as a required course and incorporating neural networks and LLMs into course material.
- Assisted graduate students in finding funding through grants or teaching assistant opportunities on campus.
- Chaired the Master's Admissions committee and coordinated with the Director of JPSM on recruitment efforts.

Joint Program in Survey Methodology, University of Maryland, Associate Research Professor — College Park, MD — July 2024 – present — 1 year 9 months

- Developed and implemented innovative teaching assignments and tools including: cloud-based programming workbooks using JupyterHub (multiple classes); a board game decision analysis project for an undergraduate class (INST 354: Decision Making for Information Science); a data disclosure-proofing project (SURV 622: Fundamentals of Data Collection); and a custom R package containing tutorials (SURV 613: Machine Learning for Social Science).
- Taught classes on introductory statistics, data science, machine learning, and programming for both undergraduate and graduate students, utilizing traditional and flipped classroom formats in online, in-person, and HyFlex modes.
- As part of an NSF grant, managed and oversaw the development and day-to-day operations of data analytics training courses for working professionals and graduate students at the National Center for Science and Engineering Statistics (NCSES) at the National Science Foundation (NSF).

Joint Program in Survey Methodology, University of Maryland, Assistant Research Professor — College Park, MD — Aug 2020 – July 2024 — 4 years

Joint Program in Survey Methodology, University of Maryland, Lecturer — College Park, MD — Oct 2017 – Aug 2020 — 2 years 11 months

The Coleridge Initiative, Consultant/Facilitator — Sept 2020 – Jan 2022 — 1 year 5 months

- Provided guidance as a facilitator with the Applied Data Analytics program for teaching record linkage, text analysis, network analysis, and machine learning to working professionals in public policy.
- Consulted on a research projects using machine learning for imputation with large scale transaction-level purchasing data.
- Developed modularized lessons and notebooks for introductory statistics and programming in Python, R and SQL.

UCLA Institute for Digital Research and Education, Statistical Consultant — Los Angeles, CA — May 2014 – June 2016 — 2 years 2 months

- Assisted graduate students and professors in Psychology, Economics, Education, and other fields in cleaning, merging, and preparing data for analysis.
- Helped clients with their dissertations, publications, and other projects using methods such as mixed modeling, principal components analysis, simulation, and more using a variety of statistical software.

UCLA Institute for Pure and Applied Mathematics, Academic Mentor — UCLA — 2015 – 2017 — 2 years

- Mentored a team of four undergraduate students with a research project for the Research in Industrial Projects for Students (RIPS) program in which an industry sponsor provides a real problem of interest.
- Provided a cooperative team environment for research by encouraging collaboration between students and advising students throughout the full research lifecycle.

Publications

Using a Stopping Rule to Optimize Cost-Quality Tradeoffs in a Large, Mixed-Mode Survey: A Simulation Study
James Wagner, Brady West, **Brian Kim**, Deji Suolang, Curtiss Engstrom, Jennifer Sinibaldi
www.doi.org/10.1177/0282423X241287452

Hidden Population Size Estimation and Diagnostics Using Two Respondent-Driven Samples with Applications in Armenia
B. Kim, L. G. Johnston, T. Grigoryan, A. Papoyan, S. Grigoryan, K. R. McLaughlin
www.doi.org/10.1002/bimj.202200136

Using Geographical Data and Rolling Statistics for Diagnostics of Respondent-Driven Sampling
B. Kim, M. Ogwal, E. Sande, H. Kiyingi, D. Serwadda, W. Hladik
www.doi.org/10.1016/j.socnet.2020.05.001 (Social Networks. 69, 74-83)

Population Size Estimation Using Multiple Respondent-Driven Sampling Surveys
B. Kim, M. Handcock
www.doi.org/10.1093/jssam/smz055 (Journal of Survey Statistics and Methodology. 9(1), 94-120)

Easy-to-Use Cloud Computing for Teaching Data Science
B. Kim, G. Henke
www.doi.org/10.1080/10691898.2020.1860726 (Journal of Statistics Education. 29:sup1, S103-S111)

Partnering with a global platform to inform research and public policy making
F. Kreuter, N. Barkay, A. Bilinski, A. Bradford, S. Chiu, R. Eliat, J. Fan, T. Galili, D. Haimovich, *B. Kim*, S. LaRocca, Y. Li, K. Morris, S. Presser, T. Sarig, J. A. Salomon, K. Stewart, E. A. Stuart, R. Tibshirani
www.doi.org/10.18148/srm/2020.v14i2.7761 (Survey Research Methods, 14(2), 159-163)

Software and Data

sspse: Estimating Hidden Population Size using Respondent Driven Sampling Data
M. Handcock, K. Gile, *B. Kim*, K. McLaughlin
cran.r-project.org/package=sspse

The University of Maryland social data science center global COVID-19 trends and impact survey, in partnership with Facebook
J. Fan, Y. Li, K. Stewart, A. R. Kommareddy, A. Garcia, J. O'Brien, A. Bradford, X. Deng, S. Chiu, F. Kreuter, N. Barkay, A. Bilinski, *B. Kim*, T. Galili, D. Haimovich, S. LaRocca, S. Presser, K. Morris, J. Salomon, D. Vannette
covidmap.umd.edu/api.html

Book Chapters and Non-Refereed Publications

The Value of Science: Special Theme
J. Lane, *B. Kim*, F. Kreuter, A. Nunez
10.1162/99608f92.4a01213d (Harvard Data Science Review)

Scaling Up Data Science for the Social Sciences
B. Kim
10.1162/99608f92.d3f14ea4 (Harvard Data Science Review)

Workbooks
B. Kim, C. Kern, J. Morgan, C. Hunter, A. Kumar
textbook.coleridgeinitiative.org/chap-workbooks.html (Big Data and Social Science, 2nd Edition. Chapman and Hall/CRC Press. 333-339)

Selected Conference Presentations

A Hybrid Household and Respondent-Driven Sampling Strategy for Reaching Underrepresented Groups — Joint Statistical Meetings, Toronto, CA, August 2023

Machine Learning Model Selection for Complex Sample Survey Data — Symposium on Data Science and Statistics, August 2020

Data 8 Adoption at University of Maryland — National Workshop on Data Science Education, Berkeley, CA, June 2020

Population Size Estimation Using Multiple Respondent-Driven Sampling Surveys — Joint Statistical Meetings, Denver, CO, August 2019

Teaching Data Science Using Jupyter Notebooks and Binder — Symposium on Data Science and Statistics, Bellevue, WA, May 2019

Assessing Respondent-Driven Sampling Using Geographical Data — North American Social Network Conference, Washington, D.C., November 2018

Grants

Advancing Computational Expertise in Neuroscience and Psychology (co-PI) — May 2025 – May 2026
UMD Artificial Intelligence Interdisciplinary Institute: Course Development Grant
- Funding: $10,000

UMD Teaching and Learning Transformation Center: Experiential Learning Grant — September 2022 – August 2023
Cloud-based Active Learning for BSOS Statistics and Data Science Courses (PI)
- Funding: $135,573.72

Dean's Research Initiative — June 2022 – May 2023
An Evaluation of Sampling Methodologies for Reaching Underrepresented Groups (PI)
- Funding: $19,999

Global Privacy Monitor (GMP): Learning from External Privacy Stakeholders — June 2021 – May 2024
Awarded by Facebook, Inc.
- Funding: $2,994,731

National Science Foundation, NCSES BAA — September 2020 – August 2021
Modernizing NCSES Data Collection Approaches (co-PI)
- Funding: $88,802

National Science Foundation, ECR: 1956114 — May 2020 – May 2025
Collaborative Research: Impacts of Hard/Soft Skills on STEM Workforce Trajectories (co-PI)
- Funding: $401,716

National Agricultural Statistics Service — August 2019 – February 2024
Linked Dataset and Virtual Research Data Center (Co-PI)
- Funding: $653,490

UMD Teaching Innovations Grant June 2020 – August 2020

- Funding: $18,000

UMD Teaching Innovations Grant June 2020 – August 2020

- Funding: $5,000

UMD Year of Data Science July 2019 – May 2020

Developing a New Course in Data Science: Introduction to Data Science for Social Sciences (PI)

- Funding: $17,500

Governor's Office of Crime Control and Prevention December 2018 – July 2019

Feasibility Testing of a Respondent-Driven Sampling Approach to Recruit Human Trafficking Victims

- Funding: $30,000

Awards

- Dissertation Year Fellowship (September 2016 – June 2017)
- Collegium of University Teaching Fellows (September 2016 – June 2017)
- Graduate Dean's Scholar Award (June 2012 – September 2014)

Other Academic Work

Guest Editor Spring 2022

Harvard Data Science Review, Special Theme: Value of Science. Issue 4.2.

Reviewer

Journal of Statistics and Data Science Education

Reviewer

Journal of Survey Statistics and Methodology

Reviewer

Journal of Statistical Software

Reviewer

Public Opinion Quarterly

Reviewer

Harvard Data Science Review

Reviewer

Computational Statistics and Data Analysis

Reviewer

Journal of Official Statistics

Teaching

University of Maryland, Instructor — Oct 2017 – present

- SDSB180 (formerly BSOS180): Introduction to Data Management (Fall 2024, Spring 2025, Fall 2025, Spring 2026)
- SDSB123 (formerly BSOS188C): Social Data Science: Pathways and Applications (Fall 2024, Spring 2025, Fall 2025, Spring 2026)
- Undergraduate Courses: SDSB326 (formerly BSOS326): Python Programming for Social Science (Spring 2024)
- Undergraduate Courses: SDSB233 (formerly BSOS233): Data Science for Social Sciences (Fall 2019, Spring 2020, Spring 2022, Fall 2022, Spring 2023)
- Undergraduate Courses: INST 314: Statistics for Information Science (Fall 2018)
- Undergraduate Courses: INST 354: Decision Making for Information Science (Spring 2019)
- Graduate Courses: SURV 613: Machine Learning for Social Sciences (Fall 2020, Fall 2021, Spring 2023, Spring 2024, Spring 2025, Spring 2026)
- Graduate Courses: SURV 622: Fundamentals of Data Collection (Spring 2020, Spring 2021, Spring 2022)
- Graduate Courses: SURV 673: Introduction to Python & SQL (Fall 2018, Spring 2019, Summer 2019, Fall 2019, Fall 2020, Summer 2020, Summer 2021)
- Graduate Courses: SURV 699M: Review of Statistical Concepts (Summer 2018, Summer 2019, Summer 2020, Summer 2021)
- Graduate Courses: SURV 701: Analysis of Complex Sample Survey Data (Spring 2021)
- Graduate Courses: SURV 727: Fundamentals of Computing and Data Display (Fall 2022, Fall 2023, Fall 2024, Fall 2025)

University of California, Los Angeles, Instructor — Jan 2017 – Mar 2017

- Stat 98T: Six Degrees of Separation: Studying the World Through Social Networks (Winter 2017)

University of California, Los Angeles, Teaching Assistant — Sept 2013 – June 2016

- Stat 10: Introduction to Statistical Reasoning (Fall 2013, Winter 2014, Spring 2014, Fall 2014, Winter 2015, Spring 2015, Fall 2015, Winter 2016, Spring 2016)

Mentorship

Doctoral Students Advisor

- Harold Gomes (in progress)

Doctoral Committee Member

- Yuting Chen, University of Maryland, 2025
- Ai Rene Ong, University of Michigan, 2022
- Jordan Epistola, University of Maryland, 2022

Graduate Students on Funded Projects Supervisor

- Ujjayini Das (September 2022 – August 2023)
- Joseph Hoskisson (September 2022 – August 2023)
- Franklin Yang (September 2021 – August 2022)
- Sofi Sinozich (September 2020 – August 2021)

Service

- Faculty Coordinator, Joint Program in Survey Methodology Short Course Program
- Co-Chair, Social Data Science Program Committee
- Co-Chair, Department of Psychology Lecturer Search Committee
- Chair, Social Data Science Curriculum Committee
- Chair, Search Committee for Executive Director of the Center for Advances in Data and Measurement
- Member/Chair, JPSM/MPSM Masters Admissions Committee
- Member, School of Information Studies Lecturer Search Committee
- Member, Social Data Science Advisor Search Committee
- Member, Social Data Science Steering Committee
- Member, JPSM/MPSM Teaching Assignment Committee

Technical Strengths

Statistical Software: Highly skilled in R, SQL and Python, including various packages such as Rcpp, caret, tidyverse, tidymodels, and more (in R); and pandas, scikit-learn, nltk, tensorflow, and more (in Python).

Statistics: Highly skilled in a variety of methods, including but not limited to hypothesis testing, regression, mixed models, clustering analysis, social network analysis, social network models, network sampling methods, Monte Carlo simulation, Bayesian models, Machine Learning, text analysis, web scraping, and more.

Other Software: Skilled or proficient in the use of many other programs, including, but not limited to, Microsoft Word, Excel and Powerpoint; LaTeX.