Albert Y. Kim, PhD

CONTACT
INFORMATION

Mathematics and Statistics Department
Amherst College
202 Seeley Mudd Building
31 Quadrangle
Amherst MA, 01002

Webpage: <http://rudeboybert.rbind.io/>
Work E-mail: <aykim@amherst.edu>
Personal E-mail: <albert.ys.kim@gmail.com>
GitHub: <https://github.com/rudeboybert>
Twitter: <https://twitter.com/rudeboybert>

EDUCATION

University of Washington, Seattle WA   August 2011

Ph.D., Statistics

- Thesis Topic: *A Bayesian Surveillance System for Detecting Clusters of Non-Infectious Diseases.*
- Advisor: Professor Jon C. Wakefield.
- Area of Study: Spatial Epidemiology, Biostatistics, Markov Chain Monte Carlo Methods.

McGill University, Montreal QC   December 2003

B.Sc., Joint Honours in Mathematics and Computer Science

- *First Class Honours.*
- Minor in Management.

ACADEMIC
EXPERIENCE

Amherst College, Amherst MA   August 2017 to Present

Lecturer in Statistics, Mathematics and Statistics Department

- Courses taught:
  - Introduction to Statistics via Modeling (STAT 135): Introductory statistics and data science course.
  - Advanced Data Analysis (STAT 495): Senior capstone course on machine learning in preparation for senior theses

Middlebury College, Middlebury VT   August 2015 to July 2017

Assistant Professor of Statistics, Mathematics Department

- Courses taught:
  - Introduction to Statistical and Data Sciences (MATH 116): Introductory statistics aimed primarily at non-math majors.
  - Introduction to Data Science (MATH 216, new course): Analysing large and complex data sets using advanced statistial software.
  - Statistical Learning (MATH 218, new course Spring 2017): Machine/statistical learning at a level appropriate for a student who hasn't taken an upper-level undergraduate probability course.
  - Theory of Statistics (MATH 311): Upper-level undergraduate statistics course.

Reed College, Portland OR   August 2013 to July 2015

Visiting Assistant Professor of Statistics, Mathematics Department

- Responsibilities:
  - Teaching undergraduate probability and statistics courses.
  - Supervising senior theses.
- Courses taught:
  - Introduction to Probability and Statistics (MATH 141): Introductory statistics aimed primarily at non-math majors.
  - Probability (MATH 391): Upper-level undergraduate probability course.
  - Mathematical Statistics (MATH 392): Upper-level undergraduate statistics course.
  - Case Studies in Statistical Analysis (MATH 241, new course): Analysing large and complex data sets using advanced statistial software.

University of Washington Seattle, WA   June 2010 to August 2010

Pre-Doctoral Instructor, Statistics Department

- Courses taught:
  - Statistics for Engineers and Scientists (STAT/MATH 390): Introductory statistics for non-statistics science and engineering majors.

University of Washington Seattle, WA   September 2004 to June 2005

Teaching Assistant, Statistics Department

- Courses taught:
  - Basic Statistics (STAT 220): Introductory statistics for non-science majors.

PROFESSIONAL
EXPERIENCE

Decision Support Engineering Analyst   June 2011 to March 2013

Ads Metrics, Google Inc., Mountain View, CA

- Manager: Nicholas Chamandy
- Responsibilities:
  - Quantitative analyst in Ads Metrics (analyst sub-team of Google AdWords), tasked with ensuring the quality of search ads shown on google.com.
  - Extensive use of Google's internal MapReduce system for distributed computing on clusters of computers
  - Use of advanced statistical techniques for analysis in a "big data" setting.
  - Communicated results and findings to other analysts, engineers, and managers in the Ads Quality branch of Google AdWords.
  - Specific projects contributed to as an analyst: a) Long-Term Value revamping of the Google AdWords auction procedure b) Ads Human Evaluation, members of which analyze user ratings of ads shown on google.com c) Personalization effort to optimize ads targeting on google.com.

Decision Support Engineering Analyst Intern   January 2010 to March 2010

Ads Metrics, Google Inc., Mountain View, CA

- Manager: Nicholas Chamandy
- Responsibilities:
  - Same as full time position described above.
  - Specific projects contributed to as an analyst: Task Classification Project, where the intent of Google users are modeled based on search queries and ad clicks.

Peer-Reviewed Publications	[1] A.Y. Kim, C. Ismay, J. Chunn. (2018) The fivethirtyeight R Package: "Tame Data" Principles for Introductory Statistics and Data Science Courses. *Technological Innovations in Statistics Education* (Accepted)
	[2] R.D. De Veaux, M. Agarwal, M. Averett, B.S. Baumer, A. Bray, T.C. Bressoud,L. Bryant, L.Z. Cheng, A. Francis, R. Gould, A.Y. Kim, M. Kretchmar, Q. Lu, A. Moskol, D. Nolan, R. Pelayo, S. Raleigh, R.J. Sethi, M. Sondjaja, N. Tiruviluamala, P.X. Uhlig, T.M. Washington, C.L. Wesley, D. White, and P. Ye. (2017) Curriculum Guidelines for Undergraduate Programs in Data Science. *Annual Review of Statistics and Its Application*, 4(2), 1-16.
	[3] A.Y. Kim, and J. Wakefield. (2016) A Bayesian Method for Cluster Detection with Application to Brain and Breast Cancer in Puget Sound. *Epidemiology*, 27(3), 347-55.
	[4] A.Y. Kim, and A. Escobedo-Land. (2015) OkCupid Profile Data for Introductory Statistics and Data Science Courses. *Journal of Statistical Education*, 23(2).
	[5] J. Wakefield, and A.Y. Kim. (2013) A Bayesian Model for Cluster Detection. *Biostatistics*, 14(4), 752-765.
	[6] A.Y. Kim, C. Marzban, D.B. Percival, and W. Stuetzle (2009) Using Labeled Data to Evaluate Change Detectors in a Multivariate Streaming Environment. *Signal Processing*, 89(12), 2529-2536.
Preprints or In Preparation	[7] C. Ismay, A.Y. Kim. (2018) ModernDive: An Introduction to Statistical and Data Sciences via R
	[8] T. Singh, A.Y. Kim. (2018) Regression and Matching for Causal Inference.
	[9] D. Allen, A.Y. Kim. (2018) Estimating Species-Specific Competition Coefficients with a Bayesian Hierarchical Model of the Neighborhood Effect of Competition on Tree Growth.
Awards	• Dorothy M. Gilford Teaching Award: University of Washington Department of Statistics, 2005. Awarded annually to recognize the best teaching assistant. • Bill and Hilde Birnbaum Fellowship: University of Washington Department of Statistics, 2004. Scholarship awarded annually to one incoming student.
Student Advising and Research	• Senior Thesis: James Burke, "Baseball as a Markov Chain: A Bayesian Approach" Spring 2017. • Independent Study (MATH 500): Trisha Singh, reading course on causal inference, Spring 2017. • Independent Study (MATH 500): Mohamed Hussein, reading course on statistical/machine learning, Fall 2016. • Senior Thesis: Blake Rosenthal, "Mapping Oregon Groundwater: A Geo-Statistical Anaylsis in Spatial Interpolation," Spring 2015. • Senior Thesis (co-advised with Prof. B. Thomas): Carl Proepper, "Generalized statistical techniques for differentiating dark matter models at the Large Hadron Collider," Spring 2015. • Independent Study (MATH 482): Liam Bowcock, reading course on measure theoretic probability, Spring 2015. • Paper: Adriana Escobedo-Land, "OkCupid Profile Data for Introductory Statistics and Data Science Courses," 2015.

- Summer Science Research Fellowship: Jacob Menick, "Evaluating Latent Dirichlet Allocation Topics," Summer 2014.
- Senior Thesis: Kevin Gallagher, "Building a Better Mortgage-Backed Security: Correctly Pricing Associated Risks," Spring 2014.
- Senior Thesis: Tristan Hechtel, "Pay It Forward" Tuition: An Econometric Analysis," Spring 2014.
- Senior Thesis: Torrey Payne, "The Generalist Bias: Estimating the Value of Three-Point Shooting in the National Basketball Association," Spring 2014.
- Senior Thesis: Joan Wang, "Food Hinterlands: The Sprawling of Food Deserts," Spring 2014.

PRESENTATIONS AND WORKSHOPS

- Meetup, December 2017: Modeling the effects of interspecies competition on the growth of trees. 5College Stats and Data Science Research Bytes, Western Mass Statistics and Data Science Meetup, Amherst MA.
- Tech Talk, May 2016: Using GitHub for Education to Encourage Open Learning and Facilitate Feedback. Middlebury College "Digital Liberal Arts Behind the Scenes series", Middlebury VT.
- Tech Talk, May 2016: Using GitHub for Education to Encourage Open Learning and Facilitate Feedback. Middlebury College "Digital Liberal Arts Behind the Scenes series", Middlebury VT.
- Meetup, April 2016: How to Teach Data Science. Burlington Data Scientists Meetup, Burlington VT.
- Paper, August 2015: A Bayesian Model for Detecting Cluster Detection. New Researchers Conference, Joint Statistical Meetings, Seattle WA.
- Tech Talk, June 2015: Teaching Data Science to Undergrads. 729 Miles of Technology Conference, Reed College, Portland OR.
- Tech Talk, June 2015: Teaching data science to undergrads: an ex-Googler's tales from the trenches. Google Inc, Mountain View, CA.
- Paper, February 2014: A Bayesian Model for Detecting Clusters of Non-Infectious Diseases. Oregon Chapter of the American Statistical Association, Portland OR.
- Paper, August 2010: A Bayesian Model for Detecting Clusters of Non-Infectious Diseases. Joint Statistical Meetings, Vancouver BC Canada.
- Paper, August 2008: Using Labeled Data to Evaluate Change Detectors in a Multivariate Streaming Environment. Joint Statistical Meetings, Denver CO.
- Paper, May 2008: Using Labeled Data to Evaluate Change Detectors in a Multivariate Streaming Environment. Interface Conference, Durham NC.
- Report, February 2006: Changepoint Detection in Multivariate Data Streams. Counter Improvised Explosive Device Meeting, Naval Research Laboratory, Washington DC.

PANELS AND WORKSHOPS

- Career speakers panel, September 2017. StatFest, Emory University, Atlanta GA.
- "A Fully Customizable Textbook for Introductory Statistics/Data Science Courses," USCOTS, June 2017, State College PA
- "Open Source Resources for Mathematics: Benefits and Costs," Mathematical Association of America Mathfest conference, August 2014, Portland OR.
- Career speakers panel, May 2014. Statistical Society of Canada Annual Meeting Student Conference, Toronto ON.

COMPUTING SKILLS

- Programming languages: C, C++, Python, MySQL, Java.
- Applications: R, Git, MapReduce, MATLAB.
- Maintainer of the following `R` open-source software packages:
  - `SpatialEpi`: Data and methods for spatial epidemiology.

- - ○ `resampledata`: Package of data sets from "Mathematical Statistics with Resampling in R" by Chihara and Hesterberg.
    - ○ `fivethirtyeight`: Data and Code Behind the Stories and Interactives at FiveThirtyEight.
  - Contributor to the following `R` open-source software packages:
    - ○ `oilabs`: Data and code necessary for R labs accompanying OpenIntro textbook.
    - ○ `infer`: Package for tidy statistical inference

JOURNAL REVIEW
SERVICE
- The American Statistician
- Journal for Quantitative Analysis in Sports
- Technology Innovations in Statistics Education
- Journal of Statistics Education

EXTERNAL
SERVICE
- Organizing committee, 2018 StatFest: a one-day event aimed at encouraging undergraduate students from under-represented groups to consider graduate studies and careers in the statistical sciences.
- Organizing committee, 2017 ASA DataFest.
- Volunteer for OpenIntro open-source introductory statistics textbook.

COLLEGE SERVICE
- Amherst College StatFellows coordinator: StatFellows are undergraduate statistics ambassadors for Amherst College performing tutoring and statistical consulting across campus, 2017-2018
- Faculty lead for Green Chicken student math competition between Williams and Middlebury Colleges (co-lead with Prof. E. Proctor), 2016.
- Faculty lead for R Community of Practice for Middlebury Faculty, 2016-present.
- Co-organizer of inaugural ASA DataFest Vermont (with Prof. P. Yates, Saint Michael's College), 2016.
- Chief organizer of ASA DataFest Vermont (Middlebury College), 2017.
- Volunteer judge for ASA DataFest Five Colleges (UMass Amherst) 2016 and 2017.
- Co-creator of Reed College statistics concentration within mathematics major (with Prof. A. Jones), 2015.

WORKSHOPS
ATTENDED
- 2016 NES/MAA Vermont Workshop on "Teaching Calculus Now" by Prof. David Bressoud, Saint Michael's College, Colchester VT, October 2016.
- Park City Mathematics Institute Summer Session Faculty Program on establishing a curriculum for Data Science, Midway UT, July 2016.
- Big Data: Implications for the Liberal Arts Curriculum, Wellesley College, Wellesley MA, January 2016.
- Advanced R, JSM 2015, Seattle WA, July 2015.
- New Researchers Conference, JSM 2015, Seattle WA, July 2015.
- OpenStreetMaps workshop, Foss4G Free and Open Source Software for Geospatial conference, Portland OR, Sept 2014.

PERSONAL
- Languages: English, French (fluent), Korean (conversational).
- Interests: History, hiking, standup paddleboarding, hockey, skiing, bicycling, and backcountry camping.

LAST UPDATED   2018/01/17