

# ADAM KAPELNER

516-435-6795 ◊ kapelner@qc.cuny.edu ◊ (updated: October 15, 2023)

## EDUCATION

---

<b>Wharton School of the University of Pennsylvania</b> Ph.D. in Statistics advised by Abba Krieger and Edward George	May 2014
<b>Wharton School of the University of Pennsylvania</b> A.M. in Statistics advised by Dean Foster	May 2012
<b>Stanford University</b> B.S. in Mathematical & Computational Science (minors in Physics & Economics)	June 2006

## ACADEMIC EMPLOYMENT

---

<b>Queens College</b> Associate Professor of Mathematics	Aug 2021 - present
Director, Undergraduate Data Science and Statistics Program	Aug 2019 - present
Assistant Professor of Mathematics	Aug 2014 - Jul 2021
<b>The Technion, Israel Institute of Technology</b> Visiting Scholar, Faculty of Industrial Engineering & Management	Jul 2018 - present

## RESEARCH INTERESTS

---

Data Science, Machine Learning, Experimental Design, Randomization, Statistical Software, Crowdsourced Social Science Experiments, Biomedical Applications, Educational Tech

## PATENTS

---

**Systems and Methods for Treatment Selection** #11605463    Granted Mar 14, 2023  
(with Armstrong, Caitrin; Benrimoh, David; Fratila, Robert; Kleinerman, Akiva; Mehltritter, Joseph; and Rosenfeld, Ariel)

## PUBLICATIONS

---

# citations:  $\approx 3,100$ ; h-index: 17, i10-index: 22    (see Google Scholar profile)

### Statistical Theory & Methodology

- Azriel, D., Krieger, A. & Kapelner, A. The Optimality of Blocking Designs in Equally and Unequally Allocated Randomized Experiments with General Response *in review at Journal of Causal Inference, Arxiv* (link)
- Kapelner, A., Krieger, A. & Azriel, D. The Role of Pairwise Matching in Experimental Design for an Incidence Outcome *in press at Australian & New Zealand Journal of Statistics, Arxiv* (link)
- Krieger, A., Azriel, D., Sklar, M. & Kapelner, A. (2022) Improving the Power of the Randomization Test *Communications in Statistics — Theory and Methods* (link)

- Krieger, A., Azriel, D. & Kapelner, A. (2021) Better Experimental Design by Hybridizing Binary Matching with Imbalance Optimization *Canadian Journal of Statistics* (link)
- Kapelner, A., Sklar M., Krieger, A., & Azriel, D. (2021) Optimal Rerandomization via a Criterion that Provides Insurance Against Failed Experiments. *Journal of Statistical Planning and Inference* 219, 63–84 (link)
- Kapelner, A. & Krieger, A. (2021) A Matching Procedure for Sequential Experiments that Iteratively Learns which Covariates Improve Power *Biometrics* (link)
- Kapelner, A., Bleich, J., Levine, A., Cohen, Z. D., DeRubeis, R. J. & Berk, R. A. (2021) Evaluating the Effectiveness of Personalized Medicine with Software. *Frontiers in Big Data — Medicine and Public Health* 4 (8), 1–19 (link)
- Kapelner, A., Shalit, U., Krieger, A., Sklar, M. & Azriel, D. (2021) Harmonizing Optimized Designs with Classic Randomization in Experiments. *The American Statistician* 75 (2), 195–206 (link)
- Blanford, W. J., Jofat, D. & Kapelner, A. (2020). Solution Density Models as Functions of Sodium Chloride, Hydroxypropyl- $\beta$ -cyclodextrin, and Temperature (278.15–333.15 K) via Progressive Linear and Stepwise Regression. *Journal of Chemical & Engineering Data* 65 (10), 4735–4750 (link)
- Krieger, A., Azriel, D. & Kapelner, A. (2019) Nearly Random Designs with Greatly Improved Balance. *Biometrika* 106 (3), 695–70 (link)
- Kapelner, A., Krieger, A. & Blanford, W. J. (2016). Optimal Experimental Designs for Estimating Henry’s Law Constants via the Method of Phase Ratio Variation. *Journal of Chromatography A* 1468, 183–191 (link)
- Kapelner, A. & Krieger, A. (2014). Matching on-the-fly in Sequential Experiments for Higher Power and Efficiency. *Biometrics* 70 (2), 378–388 (link)
- Azriel, D., Krieger, A. & Kapelner, A. Optimal Experimental Designs for Extreme Noise and Allocations *in preparation for Scandinavian Journal of Statistics*
- Krieger, A., Azriel, D. & Kapelner, A. Greedy Pair Switching for High-Performance Equal Groups Clustering *in preparation for Biometrika*

### Machine Learning and Data Science

- Kapelner, A. & Bleich, J. (2016). `bartMachine`: Machine Learning with Bayesian Additive Regression Trees. *Journal of Statistical Software* 70 (4) (link)
- Kapelner, A. & Bleich, J. (2014). Prediction with Missing Data via Bayesian Additive Regression Trees. *Canadian Journal of Statistics* 43 (2) 224–239 (link)
- Bleich, J., Kapelner, A., George, E. I. & Jensen, S. T. (2014). Variable Selection Inference for Bayesian Additive Regression Trees. *Annals of Applied Statistics* 8 (3) 1750–1781 (link)
- Goldstein, A., Kapelner, A., Bleich, J. & Pitkin, E. (2014). Peeking Inside the Black Box: Visualizing Statistical Learning with Plots of Individual Conditional Expectation. *Journal of Computational & Graphical Statistics* 24(1), 44–65 (link)
- Bleich, J & Kapelner, A. (2014) Bayesian Additive Regression Trees With Parametric Models of Heteroskedasticity. *Arxiv* (link)
- Berk, R., Bleich, J., Kapelner, A., Henderson, J., Barnes, G., Kurtz, E. (2014) Using Regression Kernels to Forecast A Failure to Appear in Court *Arxiv* (link)

- Antonaros, P., Kapelner, A. & Hanusa, C. Modeling the Artistic Beauty of Mathematical Art via an Ensemble of Deep Learning and a Random Forest *in preparation for the Journal of Mathematics and the Arts*

### **Crowdsourcing and Social Science**

- Weinberg, D. B. & Kapelner, A. Do book consumers discriminate against Black, female, or young authors? *PLoS One* 17(6): e0267537 (link)
- Kapelner, A. & Weinberg, D. B. (2019) Do Readers Judge Books by Author Gender? Results from a Randomized Experiment. *Socius* 5 (link)
- Weinberg, D. B. & Kapelner, A. (2018) Comparing gender discrimination and inequality in indie and traditional publishing. *PLoS One* 13 (4) e0195298 (link)
- Schwartz, H. A., Eichstaedt, J., Blanco, E., Agrawal, M., Dziurzyński, L., Kern, M. L., Kapelner, A., Park, G., Jha, S., Stillwell, D., Kosinski, M. & Ungar, L. H. (2016) Predicting individual well-being through the language of social media. *Biocomputing: Proceedings of the Pacific Symposium* 516–527 (link)
- Chandler, D. & Kapelner, A. (2013) Breaking Monotony with Meaning: Motivation in Crowdsourcing Markets. *Journal of Economic Behavior & Organization*, 90: 123-133 (link)
- Kapelner, A., Kaliannan, K., Schwartz, H. A., Ungar, L. H. & Foster, D. P. (2012) New Insights from Coarse Word Sense Disambiguation in the Crowd. *CoLING* (link)
- Kapelner, A. & Chandler, D. (2010) Preventing Satisficing in Online Surveys. *Proceedings of CrowdConf* (link)

### **Education Applications**

- Adlof, A., Baron, L. S., Scoggins, J., Kapelner, A., McKeown, M. G., Perfetti, C., Miller, E., Soterwood, J. & Petscher, Y. (2019) Accelerating Adolescent Vocabulary Growth: Development of an Individualized, Web-based, Vocabulary Instruction Program. *Language, Speech, and Hearing Services in Schools* 50 (4): 579-595 (link)
- Kapelner, A., Nessaiver, S., Soterwood, J. & Adlof, A. (2018) Predicting Contextual Informativeness for Vocabulary Learning. *IEEE Transactions on Learning Technologies* 11(1) 13-26 (link)

### **Biomedical Applications**

- Benrimoh, D., Kleinerman, A., Furukawa, A. T., Reynolds III, C. F., Lenze, E. J., Karp, J., Mulsant, B., Armstrong, C., Mehlretter, J., Fratila, R., Perlman, K., Israel, S., Popescu, C., Golden, G., Qassim, S., Anacleto, A., Tanguay-Sela, M., Kapelner, A., Rosenfeld, A., Turecki, G. (2023) Towards Outcome-Driven Patient Subgroups: A Machine Learning Analysis Across Six Depression Treatment Studies. *The American Journal of Geriatric Psychiatry* (link)
- Kleinerman, A., Rosenfeld, A., Benrimoh, D., Fratila, R., Armstrong, C., Mehlretter, J., Shneider, E., Yaniv-Rosenfeld, A., Karp, J., Reynolds, C.F., Turecki, G & Kapelner, A. (2021) Treatment selection using prototyping in latent-space with application to depression treatment *PLOS ONE* 16(11): e0258400 (link)

- Benrimoh, D., Israel, S., Fratila, R., Armstrong, C., Perlman, K., Rosenfeld, A. & Kapelner, A (2021) Editorial: ML and AI Safety, Effectiveness and Explainability in Healthcare. *Frontiers in Big Data* 4, 1–54 (link)
- Bleich, J., Cole, B., Kapelner, A., Baillie, C. A., Gupta, R., Hanish, A., Calgua, E., Umscheid, C. & Berk, R. (2021) Using Random Forests with Asymmetric Costs to Predict Hospital Readmissions *medrxiv* (link)
- Mehlretter, J., Fratila, R., Benrimoh, D.A., Kapelner, A., Perlman, K., Snook, E., Israel, S., Miresco, M. & Turecki, G. (2020) Differential Treatment Benefit Prediction For Treatment Selection in Depression: A Deep Learning Analysis of STAR\*D and CO-MED Data *Computational Psychiatry* 4, 61–75 (link)
- Clarke, G. P. & Kapelner, A. (2020) The BART Formula for Safe Machine-Learning Based IOL Predictions. *Frontiers in Medicine* 3 (46) 1–10 (link)
- Schoeler, N., Bell, G., Yuen, A., Kapelner, A., Heales, S. J. R., Cross, J. H. & Sisodiya, S. (2017) Acetyl carnitine and association with response to ketogenic dietary therapies. *Epilepsia* 58 (5), 893-900 (link)
- Kapelner, A. & Vorsanger, M. (2015) Starvation of Cancer via Induced Ketogenesis and Severe Hypoglycemia. *Medical Hypotheses*, 84(3): 162–168 (link)
- Chang, A. Y., Bhattacharya, N., Mu, J., Setiadi, A. F., Carcamo-Cavazos, V., Lee, G. H.; Simons, D. L., Yadegarynia, S., Hemati, K., Kapelner, A., Zheng, M., Krag, D. N., Schwartz, E. J., Chen, D. Z. & Lee, P. P. (2013) Spatial organization of dendritic cells within tumor draining lymph nodes impacts clinical outcome in breast cancer patients. *Journal of translational medicine*, 11(1): 242 (link)
- Setiadi, A. F.; Ray, N. C., Kohrt, H. E., Kapelner, A., Carcamo-Cavazos, V., Levic, E. B., Yadegarynia, S., van der Loos, C. M., Schwartz, E. J., Holmes, S. & Lee, P. P. (2010) Quantitative, architectural analysis of immune cell subsets in tumor-draining lymph nodes from breast cancer patients and healthy lymph nodes. *PloS one*, 5(8): e12420 (link)
- Holmes, S., Kapelner, A. & Lee, P. P. (2009) An interactive java statistical image segmentation system: Gemident. *Journal of Statistical Software*, 30(10): 1–20 (link)
- Kapelner, A., Lee, P. P. & Holmes, S. (2007) An interactive statistical image segmentation and visualization system. *in proceedings of IEEE, Medivis* (link)

## GRANTS AWARDED

---

<b>Air Force Research Laboratory</b>	2022-2026
Understanding Who's Vulnerable and Why?, \$3,073,000 Principal Investigator	
<b>Israel-USA Binational Science Foundation (BSF)</b>	2019-2024
More Powerful Experiments via Harmonizing Classic Randomization with Modern Optimization, 2018112, \$167,000 Co-PI	
<b>PSC CUNY</b>	2019-2020
A Natural Field Experiment on Race and Gender Discrimination in the Gig Economy, TRADB-50-65, \$6,000	
<b>MQ Foundation</b>	May 2017 - Jul 2018
The Stratified Medicine Approaches for Treatment Selection Mental Health Prediction Tournament, \$3,200 • Winner of tournament	

**PSC CUNY** 2017-2018  
 Optimal Experimentation: Trading Randomization for Balance, TRADA-48-469, \$3,500  
**MQ's Psy-IMPACT** Nov 2014 - Dec 2018  
 • Served as an expert consultant and investigator to the grant  
**U.S Dept. of Education, Inst. of Educational Science** Sept 2013 - Aug 2018  
 Grant #R305A130467, \$1,500,000, PI: Suzanne Adlof, University of South Carolina. I  
 cowrote and served as the expert consultant to the grant.

## GRANTS APPLIED

---

**ONR N00014-20-S-F005** (applied 2020), \$750K Co-PI 2021 - 2024  
 Data Science as Pathway to Naval Careers for Diverse Students at CUNY Queens College  
**NSF PD18-1269** (applied 2018), \$500K Co-PI 2019 - 2023  
 More Powerful Experiments via Harmonizing Classic Randomization with Optimization  
**PSC CUNY** (applied 2015) TRADA-47-330, \$2500 2016  
 Demonstrations of Inference for Personalized Medicine Treatment Models  
**NSF Environmental Engineering program PD 20-1440**, \$422K 2023 - 2026  
 Impact, Analysis, and Advanced Modeling of Miscible Displacement Transport Studies of Monoaromatic Compounds in Sandstone Cores

## HONORS AND AWARDS

---

President's Award for Excellence in Teaching Mar 2023  
 Certificate for Highly Cited Research in the Journal  
 of Economic Behavior and Organization Jan 2017  
 National Science Foundation Graduate Research Fellowship May 2010 - April 2013  
 J. Parker Bursk Memorial Award for Excellence in Research Dec 2013  
 Donald S. Murray Award for Excellence in Teaching Dec 2012  
 Intel Science Talent Search Semifinalist Jan 2002

## FORMAL PRESENTATIONS AND TALKS

---

Kapelner A., The Role of Pairwise Matching in Experimental Design for an Incidence Outcome, Queens College, CSCI Seminar Sept 2022  
 Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Harvard University, Applied Statistics Lecture Series Feb 2020  
 Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Design & Analysis of Experiments Conference Oct 2019  
 Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Atlantic Causal Inference Conference May 2019  
 Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Wharton Statistics Faculty Seminar Mar 2019  
 Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Economics Dept Seminar, Queens College Feb 2019  
 Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, *EMR-IBS, '18, Jerusalem, Israel* Dec 2018  
 Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Haifa University Statistics Faculty Seminar Oct 2018

- Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Tel Aviv University Statistics Faculty Seminar Oct 2018
- Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Hebrew University of Jerusalem Statistics Faculty Seminar Oct 2018
- Kapelner A., Harmonizing Fully Optimal Designs with Classic Randomization in Experiments, Weizmann Institute of Science Statistics Faculty Seminar Oct 2018
- Kapelner A., Personalized Medicine Inference via the R Package PTE, Industrial Engineering and Management Seminar, the Technion July 2018
- Kapelner A., Personalized Medicine Modeling with Survival and Incidence Endpoints *TSIL, '18, London, England* Jun, 2018
- Kapelner A., Personalized Medicine Inference and Machine Learning via the R Packages PTE & YARF *TSIL 1/2 Day Workshop, '18, London, England* Jun, 2018
- Kapelner, A., Weighted Matching on-the-fly: Improved Sequential Allocation with Higher Power and Efficiency *SAE '18 Shanghai, China* Jun, 2018
- Kapelner A., Weighted Matching on-the-fly: Improved Sequential Allocation with Higher Power and Efficiency, Technical Seminar, Amazon Inc. June 2018
- Kapelner, A., Starving Cancer through Induced Ketogenesis *QC Biology Symposium* Jan, 2018
- Kapelner, A., YARF: A Fully Customizable Non-Parametric Regression Toolbox *The Technion, seminar series* Jul, 2017
- Kapelner, A., Weighted Matching on-the-fly: Improved Sequential Allocation with Higher Power and Efficiency *The 6th International Workshop in Sequential Methodologies (IWSM '17 Rouen, France)* Jun, 2017
- Kapelner, A., Predicting Contextual Informativeness for Vocabulary Learning *Kasisto, Inc. seminar series* Nov, 2016
- Kapelner, A., Optimal experimental designs for estimating Henry's law constants via the phase ratio method *ACS National Meeting* Aug, 2016
- Clarke, G., Hill, W., Kapelner, A. Data-Driven IOL Calculations *Amer. Society of Cataract and Refractive Surgeons Annual Meeting* May, 2016
- Kapelner, A., Inference for Personalized Medicine Models *UPenn Treatment Lab* June, 2016
- Kapelner, A., Better Experiments on MTurk? NYU Statistics Seminar May, 2016
- Jensen, S., Kapelner, A., Variable Selection with Bayesian Additive Regression Trees. ENAR Mar, 2015
- Kapelner, A., Experiments via Crowdsourcing: A New Platform for Social Science Research? Economics Dept Seminar, Queens College Feb 2015
- Kapelner, A., Better Randomization via Greedy Pair Switching *IMS China* Jul, 2015
- Kapelner, A., Ungar L. Crowdsourcing for Statisticians. *JSM*, Continuing Education Course Aug 2013
- Kapelner, A., Chandler D. Preventing Satisficing in Online Surveys. *CrowdConf* Oct 2010

## INFORMAL PRESENTATIONS AND TALKS

---

- Episode 109 - Experimental Design *Local Max (Data Science Podcast)* Mar 9, 2020
- The Data Science & Statistics Mathematics Major Option at QC *NYC Tech Talent Pipeline*

*Conference*  
Predicting Informativeness from Context *QC Math Club*

May 14, 2018  
Feb 5, 2018

## TEACHING EXPERIENCE

---

### Queens College, City University of New York

Math 342W/642/RM742 (Data Science / Machine Learning Fund., 6cr) Jan 2018 - present  
Math 340/640 (Probability Theory for Data Science, 4cr) sch. for Aug 2023  
Math 341/641 (Statistical Theory for Data Science, 4cr) sch. for Aug 2023  
Math 343/643 (Computational Statistics for Data Science, 3cr) sch. for Jan 2024  
Math 369/690.3 (Statistical Inference, 3cr) Aug 2020 - Dec 2021  
Math 368/621 (Advanced Probability, 3cr) Aug 2017 - Dec 2021  
Math 341/650.3 (Bayesian Modeling, 3cr) Feb 2016 - May 2022  
Math 241 (Probability and Statistics, 3cr) Aug 2014 - Dec 2021

### The Wharton School of the University of Pennsylvania

Stat 422/722 (Predictive Analytics) Jan 2017 - Feb 2017  
Stat 101 (Probability and Statistics) May 2011 - July 2011  
Teaching Assistant for Stat 101 (Probability and  
Statistics) and Statistics 102 (Linear Regression) Sept 2009 - June 2010  
Teaching Assistant for Stat 613 (Required  
Statistics course for MBA students) Sept 2013 - Dec 2013  
Teaching Assistant for Stat 112 (Statistical Inference) Jan 2014 - Jun 2014

## NEW COURSES DEVELOPED

---

### Queens College, City University of New York

Math 340/640 (Probability Theory for Data Science, 4cr) sch. for Aug 2023  
Math 341/641 (Statistical Theory for Data Science, 4cr) sch. for Aug 2023  
Math 343/643 (Computational Statistics for Data Science, 3cr) sch. for Jan 2024  
Math 369 (Statistical Inference, 3cr) Sept 2020 - present  
Math 368 (Advanced Probability, 3cr) Sept 2019 - present  
Math 342W/642/742 (Data Science / Machine Learning Fund., 6cr) Jan 2018 - present  
Math 341 (Bayesian Modeling) Feb 2016 - present

## UNDERGRADUATE PROGRAMS DEVELOPED

---

### Queens College, City University of New York

Data Science & Statistics Option for the BA in Mathematics  
(with Chris Hanusa and Alan Sultan) Sep 2018 - present  
Actuarial Studies Track in the MA in Risk Management  
(with Cara Marshall) Jan 2018 - present  
Data Science & Statistics Speaker Series Feb 2019 - present

## GRADUATE STUDENT MENTORING EXPERIENCE

---

### **Studies in Predicting Monotonic Breakout Curves using Machine Learning**

Kennly Weerasinghe (Masters thesis)

Spring, 2021 - Summer, 2022

## UNDERGRADUATE STUDENT MENTORING EXPERIENCE

---

### **Using Convolutional Neural Nets for Predicting Digital Art Quality**

Peter Antonaros

Summer, 2022 - present

### **Studies in Applications of Predictive Analytics to Social Media Phenomena**

Namisha Singh (Honors thesis)

Winter, 2021 - Spring, 2022

### **Studies in Missing Data in Random Forests**

Abhinav Patil

Summer, 2020 - Fall, 2021

### **An Algorithm for Automatic Convergence of Random Forests**

Rebecca Horowitz

Summer, 2020

### **Studies in Optimal Experimental Design**

Abhinav Patil

Spring, 2020

### **A Nonparametric Bayesian Model for Extreme Events**

Bracha Blau

Fall, 2018

### **Studies in Fully Customizable Tree Models**

Ashok Rao

Spring, 2018

### **Studies in Optimal Experimental Design**

Bracha Blau

Spring, 2018

### **A Nonparametric Bayesian Model for Extreme Events**

Evangeline Spzylka

Summer, 2017 - Spring, 2018

### **Personalized Medicine Models for Survival**

Alina Levine

Summer, 2016 - Spring, 2018

and Xin Ling Luang (local HS student)

Summer, 2016

### **Deep Learning for Image Segmentation**

Christian Colon and Stefan Hernandez

Spring, 2016 - Spring, 2017

### **Predicting Congressmens' Party Affiliation**

Savvas Tjortjoglou

Spring, 2016

### **Web Application Engineering for Social Science Experimentation**

Rikki Katz (Honors thesis)

Spring, 2015

### **Studies in Real Analysis**

Elliot Gangaram

Spring, 2015



## DEPARTMENTAL SERVICE

---

<b>Position</b>	<b>dates</b>
Data Science & Statistics Option Advisor	9/19-present
Coaching and Interviewing Adjunct Instructors	9/19-present
Hiring Committee	9/17-present
Mathematics Minor Advisor	1/17-present
Mathematics Curriculum Committee	6/16-present
Learning Outcomes Committee	1/16-present
Faculty Development Committee	1/16-present

## COLLEGE SERVICE

---

<b>Position</b>	<b>dates</b>
Academic Senate	6/16-3/23
CUNY Office of Research - Faculty Advisory Council	12/22-present

## PROFESSIONAL SERVICE (REVIEW WORK)

---

(see Web of Science profile)

<b>Journal (or Organization)</b>	<b>Year</b>
Journal of the American Statistical Association	2023
Statistics in Medicine	2023
Journal of the American Statistical Association	2022
Statistical Science	2022
Statistics in Medicine	2022
Computational Statistics & Data Analysis	2022
Journal of the American Statistical Association	2021
PSC CUNY Grant for Statistics Research	2021
Computational Statistics & Data Analysis	2021
Journal of Agricultural, Biological, and Environmental Statistics	2021
Open Statistics	2020
INFORMS Journal on Computing	2020
International Statistical Review	2020
Biometrika	2019
Journal of the American Statistical Association	2019
PSC CUNY Grant for Statistics Research	2019
Annals of Applied Statistics	2019
PLoS One	2019
Journal of Educational & Behavioral Statistics	2018
Neural Networks	2018
IEEE Access	2018
PLoS One	2018
The R Journal	2018
PSC CUNY Grant for Statistics Research	2018
Journal of the American Statistical Association	2017
Statistical Analysis and Data Mining	2017

PSC CUNY Grant for Statistics Research	2017
PLoS One	2016
The R Journal	2016
Statistical Analysis and Data Mining	2016
International Journal of Approximate Reasoning	2016
Journal of the Royal Statistical Society (Series B)	2015
Statistics in Medicine	2015
Bayesian Analysis	2014
Medical Hypothesis	2014
Journal of Cancer Research & Therapy	2014
Annals of Applied Statistics	2013
Transactions of the Association for Computational Linguistics	2013
National Science Foundation	2010

## INDUSTRY EXPERIENCE

---

- Data Science Private Consulting** June 2014 - Present
- Prediction modeling, data mining, statistical testing for a variety of clients from tech to real estate to finance e.g. Tesorio, Rubinstein Partners, Obsidian Insurance
- Coatue, Quant Fund** May 2019 - August 2019  
*Data Scientist* *New York, NY*
- Model (or “signal”) creation for algorithm trading, optimization of trading algorithm
- DictionarySquared, Inc.** April 2010 - Dec 2018  
*Founder & CTO* *San Francisco, CA*
- Conceived and engineered a web app that teaches vocabulary via contextual snippets
  - Wharton Business Plan Semifinalist Winner
  - Applied and received federal grant money for research (see Grants section)
- Eventbrite, Inc.** April 2007- Aug 2007  
*Software Engineer* *San Francisco, CA*
- First engineer. Designed and engineered portions of their web platform.
- Stanford University, Lab of Peter Lee** June 2005 - Mar 2007  
*Staff Scientist* *Stanford, CA*
- Conceived and engineered software that finds objects in images, used to find cells in microscopic images. Uses Java-R programming, Random Forests and image processing

## OPEN SOURCE SOFTWARE ON GITHUB

---

60 ★ 25 📄	R Package: bartMachine	A flexible statistical learning suite 📄 243K
34 ★ 3 📄	R Package: ICEbox	Visualization that explain how ML models work 📄 68K
6 ★ 4 📄	Java JAR: GemIdent	Finds objects in images using machine learning
1 ★	R Package: PTE	Inference for personalized medicine models 📄 32K
	Java JAR: GemVident	Finds objects in videos using machine learning
2 📄	R Package: GreedyExperimentalDesign	Better experimental designs 📄 27K

- 1 ★ R Package: CovBalAndRandExpDes      Optimal rerandomization designs
- R Package: SeqExpMatch    High-powered matched experimental designs 📄 5K
- 3 📄 R Package: YARF      A highly customizable predictive modeling suite
- R Package: optDesignSlopeInt      Optimal designs for slope-to-intercept
- 1 ★ 1 📄 R Package: fastLogisticRegressionWrap    Fast Inference for Logistic Regression

## OTHER SOFTWARE

---

dictionaryresearch.sc.edu      Teaches high school students vocabulary

gradesly.com      Gives students transparency, helps professors administer courses

fireplacetorah.com      Automatic student pair-matching for learning texts

## TECHNICAL STRENGTHS

---

**Software Languages**    R, Java, Ruby on Rails, C++, Python, HTML/CSS/JS

**R Skills**      ggplot2, dplyr, data.table, mlr3, rJava, Rcpp, Eigen  
most ML packages

**Other Skills**      git, postgresSQL, MySQL, Linux, Gurobi, Stan, twilio messaging  
and voice APIs, grid and cloud computing, simulation, AWS