

# Philip Waggoner

YouGov America  
432 Park Avenue South, Floor 5  
New York, NY, 10016

[philip.waggoner@yougov.com](mailto:philip.waggoner@yougov.com) // [pdw2119@columbia.edu](mailto:pdw2119@columbia.edu)  
[pdwaggoner.github.io](https://pdwaggoner.github.io)

## Work

---

### Current Appointments

2021 – current Director of Data Science, *YouGov America*  
2020 – current Research Scholar, *Institute for Social and Economic Research and Policy, Columbia University*  
2023 – current Adjunct Professor of Analytics & Public Policy, *Johns Hopkins University*  
2022 – current Adjunct Professor of Data Science, *Northwestern University*  
2022 – current Adjunct Professor of Data Science, *University of Virginia*

### Past Appointments

2021 Data Scientist, *Edison Software* (acquired by *Yipit Data*)  
2019 – 2021 Assistant Instructional Professor of Computational Social Science, *University of Chicago*  
2019 Lecturer in Computational Social Science, *University of Chicago*  
2018 – 2019 Visiting Assistant Professor of Government, *College of William & Mary*  
2018 – 2019 Faculty Affiliate, *Social Science Research Methods Center, College of William & Mary*

## Editorial

---

2022 – current Co-Guest Editor (Invited), *Mathematics*, “Advances in Statistical Computing” Special Issue  
2022 – current Associate Editor, *Journal of Social Computing*  
2020 – current Editorial Board Member, *Journal of Mathematical Sociology*  
2019 – 2022 Associate Editor, *Journal of Open Research Software*

## Education & Accreditation

---

2019 Graduate Statistician (GStat), American Statistical Association

2018 Ph.D., Political Science, University of Houston  
*Committee:* Justin H. Kirkland (Chair), Brandon Rottinghaus, Jennifer H. Clark, Boris Shor,  
Justin Grimmer (Stanford University)

2017 M.A., Political Science, University of Houston

2013 M.P.A., Public Administration, School of Public Affairs, University of Colorado, Denver  
*Thesis Advisor:* Paul Teske

2011 B.A., Political Science, Colorado State University  
*Major Honors*

## Books

---

3. **Modern Dimension Reduction.** 2021. *Cambridge University Press*

2. **Introduction to R for Social Scientists: A Tidy Programming Approach** (with Ryan Kennedy). 2021. *Chapman & Hall/CRC Press*
1. **Unsupervised Machine Learning for Clustering in Political and Social Research**. 2020. *Cambridge University Press*

## Peer-Reviewed Papers

---

28. **A Batch Process for High Dimensional Imputation**. Forthcoming, *Computational Statistics*
27. **Trust in Public Policy Algorithms** (with Ryan Kennedy & Matthew Ward). 2022. *Journal of Politics*, 84(2)
26. **The Role of Personality in Trust in Public Policy Automation** (with Ryan Kennedy). 2022. *Journal of Behavioral Data Science*, 2(1)
25. **Pursuing Open-Source Development of Predictive Algorithms: The Case of Criminal Sentencing Algorithms** (with Alec MacMillen). 2022. *Journal of Computational Social Science*, 5, 89–109
24. **Uncovering the Online Social Structure Surrounding COVID-19** (with Robert Y. Shapiro, Sam Frederick, and Ming Gong). 2021. *Journal of Social Computing*, 2(2)
23. **Community Detection in Google Searches Related to 'Coronavirus'**. 2021. *Journal of Data Science*, 19(2)
22. **see: An R Package for Visualizing Statistical Models** (with Daniel Lüdecke, Mattan S. Ben-Shachar, Indrajeet Patil, Brenton M. Wiernik, and Dominique Makowski). 2021. *Journal of Open Source Software*, 6(64)
21. **A Computational Exploration of the Evolution of Governmental Policy Responses to Epidemics Before and During the Era of COVID-19**. 2021. In *The Year in C-SPAN Archives Research, Vol. 7*, edited Volume from *Purdue University Press*
20. **Pandemic Policymaking**. 2021. *Journal of Social Computing*, 2(1)
19. **Applying Dimension Reduction in Modern Data Science and Quantitative Analysis**. 2021. *Software Impacts*, 8, 100075
18. **performance: An R Package for Assessment, Comparison and Testing of Statistical Models** (with Daniel Lüdecke, Mattan Ben-Shachar Indrajeet Patil, and Dominique Makowski). 2021. *Journal of Open Source Software*, 6(60)
17. **Are there Racial Disparities in Fatal Police Shootings? Exploration with Uniform Manifold Approximation and Projection**. 2021. In *the Book of Abstracts of the 9th International Conference on Complex Networks and their Applications* (COMPLEX NETWORKS)
16. **Exploring Ideological Signals from Cosponsorship** (with Carol Downes). 2020. *Journal of Mathematical Sociology*, 45(4)
15. **Measuring Media Freedom: An Item Response Theory Analysis of Existing Indicators** (with Jonathan Solis). 2020. *British Journal of Political Science*, 51(4)
14. **Exploring the Effects of Allegations of Sexual Misconduct on Political Careers** (with Andrew Creekmore). 2020. *Social Science Journal*
13. **The Shape of and Solutions to the MTurk Quality Crisis** (with Scott Clifford, Ryan Kennedy, Tyler Burleigh, Ryan Jewell, and Nick Winter). 2020. *Political Science Research and Methods*, 8(4)  
→ Popular version in *The Washington Post's Monkey Cage*
12. **Exploring and Comparing Unsupervised Clustering Algorithms** (with Marc Lavielle). 2020. *Journal of Open Research Software*, 8(21)
11. **A Simple Method for Purging Mediation Effects**. 2020. *Journal of Statistical Theory & Practice*, 14(25)
10. **Big Data and Trust in Public Policy Automation** (with Ryan Kennedy, Hayden Le, and Myriam Shiran). 2019. *Statistics, Politics, & Policy*, 10(2)

9. **insight: A Unified Interface to Access Information from Model Objects in R** (with Daniel Lüdecke & Dominique Makowski). 2019. *Journal of Open Source Software*, 4(38)
8. **Detecting Fraud in Online Surveys by Tracing, Scoring, and Visualizing IP Addresses** (with Ryan Kennedy and Scott Clifford). 2019. *Journal of Open Source Software*, 4(37)
7. **Do Constituents Influence Issue-Specific Bill Sponsorship?** 2019. *American Politics Research*, 47(4)  
→ Invited blog posts: *LSE USAPP*, *R Street Institute LegBranch*
6. **The hhi Package: Streamlined Calculation and Visualization of Herfindahl-Hirschman Index Scores.** 2018. *Journal of Open Source Software*, 3(28)
5. **Exploring Responsiveness in the U.S. House of Representatives.** 2018. *American Review of Politics*, 36(2)
4. **The Cost of Doing Business: Congressional Requests, Cost, and Allocation of Presidential Resources** (with Brandon Rottinghaus). 2018. *Political Research Quarterly*, 71(4)
3. **Informal Politics: The Role of Congressional Letters in Legislative Behavior and Careers.** 2017. *Journal of Legislative Studies*, 23(3)
2. **Assessing the Value of State Legislative Experience and Legislative Professionalism in National Election Performance, 1974-2010.** 2017. *Social Science Journal*, 54(3)
1. **Are Samples Drawn from Mechanical Turk Valid for Research on Political Ideology?** (with Scott Clifford and Ryan Jewell). 2015. *Research & Politics*, 2(4)  
→ Cited in *The Washington Post*

## Working Papers

---

1. **Applying Regularization for Feature and Theory Selection** (with Joanna Schroeder). *Under Review.*
2. **Advances in Statistical Computing: From mathematical statistics to software-based data science** (with Dominique Makowski). *In Preparation.*
3. **The Paradox of Algorithms and Blame on Public Decisionmakers** (with Ryan Kennedy and Adam Ozer). *In Preparation.*
4. **Democratic Development of Algorithms.** *In Preparation.*

## Software (Selected)

---

→ Over 127k downloads of original published software as of Aug 2022

- Active Dev** `hdImpute` Batched High Dimensional Imputation via Chained Random Forests. On CRAN & GitHub.
- Active Dev** `plotmm` Tidy Tools for Visualizing Mixture Models, with Fong Chan and Lu Zhang. On CRAN & GitHub.
- Active Dev** `easystats` An ecosystem of R packages making statistics in R intuitive and easy, with Daniel Lüdecke, Dominique Makowski, et al. Numerous packages on CRAN & GitHub.
- Stable** `amerika` An American politics-inspired color palette generator in R. On CRAN & GitHub.
- Stable** `rIP` An R Package validating IP addresses to detect fraud in online surveys, with Ryan Kennedy and Scott Clifford. On CRAN & GitHub.
- Stable** `hhi` Calculating & visualizing Herfindahl-Hirschman Index scores. On CRAN & GitHub.
- Stable** `purging` An R Package purging mediation effects from independent variables. On CRAN & GitHub.
- Stable** `mlbstats` An R Package that computes major league baseball player metrics. On CRAN & GitHub.  
→ Included in the [Sports Analytics CRAN TaskView](#)

## Invited Talks & Workshops

---

2022 R Govys Monthly Meet Up  
2022 American Statistical Association Colorado-Wyoming Chapter  
2022 PhD Student Industry Workshop, University of Houston  
2021 Edison Software Data Science Team Meeting  
2021 American Statistical Association Student Chapter, University of Georgia  
2020 Computation Workshop, University of Chicago  
2020 Government Advances in Statistical Programming (GASP) Workshop, Washington D.C.  
2020 Center for C-SPAN Scholarship & Engagement, Purdue University  
2020 Conference on Statistical Practice, American Statistical Association, Sacramento, CA  
2019 Center for Advanced Governmental Studies, Johns Hopkins University  
2019 National Center for State Courts, Williamsburg, VA  
2019 School of Public Service, Boise State University  
2018 Departments of Statistical & Data Sciences and Government, Smith College  
2018 Government Research Colloquium, College of William & Mary  
2018 Hughes Research Laboratories–METRICS & IARPA Research Meeting, Los Angeles, CA  
2017 Department of Political Science, San Francisco State University  
2017 International Republican Institute, Washington D.C.

## Honors, Awards, Fellowships, & Grants

---

2020 Micro-Metcalf Grant, The College, University of Chicago  
2018 Conference Travel Award, College of William & Mary  
2018 Hayek Fund for Scholars Grant, Institute for Humane Studies at George Mason University  
2017–2018 William A. Steiger Fund for Legislative Studies Grant, Centennial Center, American Political Science Association  
2018 Conference Travel Award, American Political Science Association  
2014–2018 Doctoral Student Tuition Fellowship, University of Houston  
2018 Prestage-Cook Travel Award, Southern Political Science Association  
2016–2017 Conference Travel Award (x2), Department of Political Science, University of Houston  
2014–2016 Presidential Endowed Fellowship (*University-wide*), University of Houston  
2014 Robert L. Lineberry Scholarship (*Inaugural Recipient*), University of Houston  
2012–2014 Cathy Shipley Best and Brightest Colorado Local Government Fellowship  
2013 Conference Travel Award, University of Colorado, Denver  
2013 Graduate Student Award to attend the Western Social Science Association Conference  
2010–2011 Pi Sigma Alpha (*National Political Science Honors Society*), Colorado State University  
Fa. 2007 President's Leadership Program, Colorado State University

## Other Writing

---

→ *Regular contributor to the American Statistical Association CCD blog and R-Bloggers. Selected posts listed here.*

April 11, 2022 "Getting Started with Quarto." *ASA CCD Portfolio Project Blog*

Mar. 29, 2022 "Indexing from 0? The Value of R and Python for Data Science." *ASA CCD Portfolio Project Blog*

Mar. 23, 2022 "Batched Imputation for High Dimensional Missing Data Problems." *R-Bloggers*

Jan. 8, 2021 "Measuring Aggregate Policy Priorities of the U.S. House of Representatives." *The Political Methodologist*, Blog of the Society for Political Methodology

July 10, 2020 "Tidy Visualization of Mixture Models in R." *R-Bloggers*

May 29, 2019 "A New Release of rIP (v1.2.0) for Detecting Fraud in Online Surveys." *R-Bloggers*

April 2, 2019 “‘The Importance of Mass Trust in Validating Algorithms in a Public Space.’ eLetter responding to ‘The accuracy, fairness, and limits of predicting recidivism’ by Julia Dressel and Hany Farid,” (with Ryan Kennedy), *Science Advances* 4(1)

Nov. 7, 2018 “How Venezuela’s Economic Crisis is Undermining Social Science Research—about Everything” (with Ryan Kennedy, Scott Clifford, Tyler Burleigh, & Ryan Jewell). *The Washington Post*, *Monkey Cage*

Aug. 22, 2018 “Advice to Young (and Old) Programmers: A Conversation with Hadley Wickham.” *R-Bloggers*

July 13, 2018 “Constituents have Minimal Influence on their Legislators’ Policy Priorities.” *London School of Economics (LSE) USAPP Politics and Policy Blog*

April 9, 2018 “A New Package (hhi) for Quick Calculation of Herfindahl-Hirschman Index scores.” *R-Bloggers*

April 3, 2018 “Do constituents influence the work of legislators?” *R Street Institute’s LegBranch Blog*

March 28, 2018 “Introducing purging: An R package for addressing mediation effects.” *R-Bloggers*

## Teaching (Selected)

---

Probability & Statistics (*Johns Hopkins University*)

Applied Statistics with R (*Northwestern University*)

Unsupervised Machine Learning (*University of Chicago*)

Introduction to Machine Learning (*University of Chicago*)

Perspectives on Computational Modeling (*University of Chicago*)

Data Mining and Data Visualization for the Social Sciences (with Ben Soltoff, *University of Chicago*)

Data Science & Politics (Senior Seminar, *College of William & Mary*)

Introduction to Computation in R (MPP Program, *College of William & Mary*)

Research Methods (*College of William & Mary*)

Learning  $\LaTeX$  Workshop (*University of Houston*)

## Conferences

---

→ Accepted to and/or presented at many conferences from 2013 to present, including: *Joint Mathematics Meeting*, *Complex Networks*, *SIAM Mathematics of Data Science Conference*, *MPSA*, *APSA*, *SPSA*, *Data Science Conference on COVID-19 (DSCC19)*, *Texas Methods Meeting*, and others. If interested, feel free to request a copy of a paper I presented in the past.

## Service

---

### Professional

**Manuscript and Proposal Reviewer** for: *American Journal of Political Science*, *American Political Science Review*, *American Politics Research*, *American Review of Politics*, *British Journal of Political Science*, C3.ai Digital Transformation Institute, Cambridge University Press, *Economic Inquiry*, *Journal of Applied Social Science*, *Journal of Behavioral Data Science*, *Journal of Data Science*, *Journal of Mathematical Sociology*, *Journal of Open Research Software*, *Journal of Open Source Software*, *Journal of Politics*, *Journal of Social Computing*, *Legislative Studies Quarterly*, *Local Government Studies*, *PLOS ONE*, *Political Behavior*, *Political Research Quarterly*, *PS: Political Science and Politics*, *Public Opinion Quarterly*, *R Journal*, *Research & Politics*, *SAGE Open*, *Social Science Journal*, *Sociological Methods and Research*

2023 (*forthcoming*) **Awards Chair** (an invited appointed officer of both sections), American Statistical Association Sections on Statistical Computing and Statistical Graphics

2023 (*forthcoming*) **Program Committee Member** (Invited), 9th International Conference on Computational Social Science (IC2S2) (Copenhagen, Denmark)

2022 – **current Member** (Invited), American Statistical Association Committee on Professional Ethics

2022 – **current Member** (Invited), American Statistical Association Committee on Career Development

2022 – **current Member**, Topical Advisory Panel for the journal *Mathematics*

2022 **Judge** (Invited), “*This is Statistics*” *Fall Data Challenge*, American Statistical Association

2022 **Program Committee Member** (Invited), 8th International Conference on Computational Social Science (IC2S2) (Chicago)

2022 **Conference Reviewer**, JuliaCon 2022

2022 **Member** (Invited), John M. Chambers Statistical Software Award Committee, Statistical Computing and Statistical Graphics Section of the American Statistical Association

2021 **Member**, John M. Chambers Statistical Software Award Committee, Statistical Computing and Statistical Graphics Section of the American Statistical Association

2018 **Panel Chair**, Annual Meeting, American Political Science Association (Boston)

2018 **Panel Chair & Discussant**, Annual Meeting, Southern Political Science Association (New Orleans)

2017 **Panel Discussant**, Annual Meeting, American Political Science Association (San Francisco)

2017 **Graduate Student Co-Coordinator**, Texas Methods Meeting (Houston)

## University of Chicago

2019–2021 **MA Thesis Primary Advisor**: Li Liu (2020), Ying Sun (2020), Ken Chen (2020), Zhaoyang Chen (2021), Yimin Li (2021), Yanwei Pan (2021), Anqi Zhou (2021)

2019–2021 **Research Assistants** (*University of Chicago, unless otherwise noted*): Clara Fong, Dhruv Baid, Cedric Elkouh, Alec MacMillen, Abhishek Pandit, Samuel Frederick (*Columbia*), Ming Gong (*Columbia*), Lynette Dang (*UCSD*), Lu Zhang (*Emory*), Hayden Le (*College of William & Mary*), Andrew Creekmore (*College of William & Mary*)

2020–2021 **Program Lead**, MA in Computational Social Science Alumni Association

2020–2021 **Member**, Admissions Committee for the MA program in Computational Social Science

## College of William & Mary

2019 **Senior Thesis Committee Member**: Samyuktha Mahadevan

2018–2019 **Faculty Mentor** (Invited): Adam Kearney and Alexander Nocks, Winners of the Undergraduate Research “Shark Tank” Competition

## Additional Research, Training, & Experience

---

**Here & there** Consultant: statistics, data science, statistical computing, course/workshop development, computational social science, machine learning, and so on. *Reach out directly for inquiries, client lists and projects, proposals, etc.*

2017–2020 Affiliate, Machine Assisted Human Decision-making Lab, University of Houston

Su. 2018 Lecturer, School of Public Affairs, University of Colorado, Denver  
2014–2018 Research & Teaching Assistant, Department of Political Science, University of Houston, Fall, Spring, & Summer  
2013–2014 Research & Teaching Assistant, School of Public Affairs, University of Colorado, Denver  
Su. 2013 Nonresident Research Assistant for Robert Y. Shapiro, Public Opinion Project, Columbia University

## Past Professional Experience

---

2014 Policy Analyst, Colorado House of Representatives  
2012–2014 Management Associate & Court Clerk, Town of Milliken, Colorado (via *Cathy Shipley Fellowship*)

## Technical Tools & Languages

---

R, shell/bash/zsh, L<sup>A</sup>T<sub>E</sub>X, Julia, Git, AWS, Python, SQL (various flavors), HTML/CSS, Qualtrics, and others

## Memberships

---

*American Statistical Association*  
*International Statistical Engineering Association*

## References

---

*Available upon request.*