

# Benjamin Williams

Department of Business Information & Analytics  
Daniels College of Business  
University of Denver  
2101 S. University Blvd. 80208

benjamin.williams@du.edu  
Office: 592 Daniels  
Phone: 303-871-3186  
www.statswithben.com

## EDUCATION

### **Ph.D. Statistical Science**

May 2019

*Southern Methodist University, Dallas, TX*

**Thesis:** *Understanding the Effects of Record Linkage on Estimators of Total when Combining a Non-Probability and Probability Sample*

**Advisor:** *S. Lynne Stokes*

### **M.S. Statistical Science**

May 2016

*Southern Methodist University*

### **B.S. Mathematics, B.A. Economics**

May 2014

*Wheaton College, Wheaton, IL,*

*Dean's List, American Physical Society Minority Scholarship Recipient*

**Graduate Coursework:** Computational Statistics, Linear Regression, Statistical Analysis I & II, Mathematical Statistics I & II, Categorical Data Analysis, Mathematical Theory of Sampling, Machine Learning, Time Series Analysis, Applied Multivariate Analysis, Bayesian Statistics, Bayesian Hierarchical Modeling, Non-Parametric Statistics, Advanced Statistical Inference, Probability Theory

## REFEREED PUBLICATIONS

Stokes, L., **Williams, B.**, McShane, R., Zalsha, S., Elliott, A. The Impact of Non-sampling Errors on Estimators of Catch from Electronic Reporting Systems. *Journal of Survey Statistics and Methodology*. Forthcoming.

**Williams, B.** (2018). Combining a Probability and a Non-Probability Sample in a Capture-Recapture Setting. *Journal of Open Source Software*, 3(28), 886, <https://doi.org/10.21105/joss.00886>.

Guldi, J. and **Williams, B.** (2018). Synthesis and Large-Scale Textual Corpora: A Nested Topic Model of Britain's Debates over Landed Property in the Nineteenth Century. *Current Research in Digital History*, 1, <https://doi.org/10.31835/crdh.2018.01>.

Purkayastha, S., Adair, H., Woodruff, A., Ryan, L. J., **Williams, B.**, James, E., Bell, K.R. Balance Testing Following Concussion: Sensitivity of Postural Sway versus Complexity Index. *Physical Medicine and Rehabilitation*, doi: <https://dx.doi.org/10.1002/pmrj.12129>.

Purkayastha, S., **Williams, B.**, Murphy, M., Lyng, S., Sabo, T., Bell, K., Reduced Heart Rate Variability and Lower Cerebral Blood Flow Associated with Poor Cognition During Recovery Following Concussion. *Autonomic Neuroscience: Basic and Clinical*, doi: <https://doi.org/10.1016/j.autneu.2019.04.004>.

## OTHER PUBLICATIONS

**Williams, B.** (2019). Understanding the Effects of Record Linkage on Estimation of Total when Combining a Big Data Source with a Probability Sample. *Proceedings Statistics Canada Symposium 2018*, Ottawa, Canada.

Khasawneh, Q., Dworak, J., Gui, P., **Williams, B.**, Elliott, A., Muthaiah, A. (2018). Real-Time Monitoring of Test Fallout Data to Quickly Identify Tester and Yield Issues in a Multi-Site Environment. *2018 IEEE 36<sup>th</sup> VLSI Test Symposium*, pp. 1 – 6, doi: 10.1109/VTS.2018.8368661.

## **PAPERS IN PROGRESS**

**Williams, B.** and Guldi, J. A Simple Automated Hierarchical Nesting Procedure for LDA Topic Models: Understanding the Relationship Between Discourses.

## **FELLOWSHIPS**

### **Southern Methodist University Dissertation Fellow**

*Southern Methodist University – Dallas, TX*

Focus: Dissertation Research

August 2018 – May 2019

### **Dedman College Interdisciplinary Institute Fellow**

*Southern Methodist University – Dallas, TX*

Focus: Big Data in the Humanities and the Relevancy of Statistics

August 2016 – May 2017

### **Dedman College Interdisciplinary Institute Fellow**

*Southern Methodist University – Dallas, TX*

Focus: Analysis of Dallas Police Officer Shootings

September 2015 – May 2016

## **TALKS AND PRESENTATIONS**

“Understanding the Effects of Record Linkage on Estimation of Total when Combining a Big Data Source with a Probability Sample,” BigSurv Conference Oct. 2018. Contributed Talk.

“Samples, Unite! Understanding the Consequences of Combining Probability and Non-Probability Samples when Linking Records is Difficult.” *Joint Statistical Meetings*, August 2018 – Vancouver, BC, Canada, Speed Presentation.

“Understanding the Relationship Between Discourses in 19th Century British Parliamentary Debates with a Novel Algorithm for Hierarchically Nesting Topic Models,” *Current Research in Digital History*, March 2018, Arlington, VA. Contributed Talk.

“Text Mining with R,” *SMU Stat Club*, March 2018 – Dallas, TX, Invited Presentation.

“Graphics in R,” *R Users Group at SMU*, November 2017 – Dallas, TX, Invited Presentation.

“Understanding the Measurement Error for Voluntary Angler Reporting Using Electronic Devices,” *American Fisheries Society Annual Meetings*, August 2017 – Tampa, FL, Invited Talk.

“Understanding the Measurement Error on Catch Estimation for Self-Reporting and Randomly Sampled Anglers in the Gulf of Mexico,” *Joint Statistical Meetings*, August 2017 – Baltimore, MD, Poster Presentation.

“Analysis of Dallas Police Shootings: A Comparison Study,” *Dedman Fellows 2015-2016 Presentation*, April 2016 – Dallas, TX.

“PTSD and Medication Adherence in Patients with Uncontrolled Hypertension,” *CSIBS Research Symposium 2013*, July 2013 – New York, NY.

## **RESEARCH EXPERIENCE**

### **Research Consultant, NOAA Funded**

*Southern Methodist University – Dallas, TX*

*Professor Lynne Stokes, PhD – Statistical Science Department*

Focus: Estimation of Total Fish Catch in Gulf of Mexico by Recreational Anglers

May 2017 – August 2018

### **Research Assistant, NSF Funded**

*Southern Methodist University – Dallas, TX*

*Professor Jo Guldi, PhD – History Department*

Focus: Text Mining of a century of British Parliamentary Debates

May 2017 – May 2018

**Summer Institute for Training in Biostatistics, NIH funded** June – July 2013  
*Mailman School of Public Health, Columbia University – New York, NY*  
Focus: Post Traumatic Stress Disorder and Medication Adherence

**Mentored Research Seminar** January 2012 – May 2012  
*Wheaton College – Wheaton, IL*  
*Professor Darcie Delzell, PhD – Mathematics Department*  
Focus: Armed Conflict and Climate in Africa

### **TEACHING EXPERIENCE**

**Text Mining with R Practicum Co-Lecturer** September 2017 – October 2017  
*Southern Methodist University – Dallas, TX*

**Statistics and Mathematics Tutor** August 2015 – May 2016  
*Learning Enhancement Center – Southern Methodist University – Dallas, TX*

**Statistics Teaching Assistant** September 2014-May 2015  
*Statistical Science Department, Southern Methodist University – Dallas, TX*

**Mathematics/Statistics Teaching Assistant** August 2012 – May 2014  
*Mathematics Department, Wheaton College – Wheaton, IL*

### **WORK EXPERIENCE**

**Assistant Director of Consulting Center** June 2016 – June 2018  
*Statistical Consulting Center, Southern Methodist University – Dallas, TX*  
*-Clients from Industry (Non-Profit and For-Profit), Academia, and Government*  
*-Projects Include: Data Analysis; Survey Creation/Implementation/Analysis; Prediction; Machine Learning; Natural Language Processing/Text Mining; Statistical Inference*

**Consultant for Lee Park Conservancy** February 2016 – May 2016  
*Lee Park – Dallas, TX*  
*-Survey Creation, Implementation, Analysis*

**Data Quality Intern** May 2015 – August 2015  
*Atmospheric Radiation Measurement, NOAA – Boulder, CO*  
*-Develop Outlier Detection Software*

**Private Mathematics Tutor** August 2011 – April 2014  
*Wheaton, IL*

**BreakAway Spring Break Mission Trip Cabinet Chair** August 2012 – May 2013  
*Wheaton College, Wheaton, IL*

**Allstate Partnerships for Educational Excellence (APEX) Team Finance Competition** January – February 2013  
*Wheaton and Northbrook, IL*

### **GRANT EXPERIENCE**

**Mississippi-Alabama Sea Grant Consortium 2016 Red Snapper Phase I** 2016 – 2018  
Significant Contributor

**Controlling Locomotion over Continuously Varying Activities for Agile Powered Prosthetic Legs** 2017  
Statistical Consultant

## **SOFTWARE**

### ***blendR***

August 2018

An R Package for Combining a Probability Sample and a Non-Probability Sample – Available on GitHub  
Peer-reviewed and published in *Journal of Open Source Software*

## **COMPUTING SKILLS**

- R (Including ggplot2, dplyr, and other tidyverse packages)
- SAS, Python, SQL, LaTeX, MALLETT, RAT-STATS, Shiny Applications
- Advanced Beginner in Java, Tableau
- Microsoft Excel, Word, PowerPoint, Google Products
- Experience using Linux computer cluster to run analyses

## **AWARDS**

- 2018 BigSurv Student Travel Award
- 2<sup>nd</sup> Place, SMU 2018 3 Minute Thesis Competition

## **MEMBERSHIPS**

- American Statistical Association
- European Survey Research Association

## **INTERESTS**

- Statistical Sampling
- Record Linkage
- Survey Methodology
- Machine Learning
- Sampling of Visualizations
- Text Mining
- Natural Language Processing
- Interdisciplinary Statistics
- Data Visualization
- Human Rights Statistics
- Sports Statistics
- Data Quality