

# Barret Schloerke

data[c("wrangler", "explorer", "visualizer")]

Purdue University  
Department of Statistics  
250 N. University St.  
West Lafayette, Indiana  
☎ (515) 290 0777  
✉ [schloerke@gmail.com](mailto:schloerke@gmail.com)  
in [schloerke](#)  
🌐 [schloerke](#)

## Education

2013  
2017 **PhD of Statistics**, *Purdue University*, West Lafayette, IN

Advisor: Dr. William Cleveland and Dr. Ryan Hafen

Anticipated graduation: December 2017

2013  
2014 **Master of Science in Mathematical Statistics**, *Purdue University*, West Lafayette, IN  
GPA: 3.69 / 4.0

2012  
2012 **Student for PhD of Statistics**, *Rice University*, Houston, TX  
Advisor: Dr. Hadley Wickham

2006  
2010 **Bachelor of Science in Computer Engineering**, *Iowa State University*, Ames, IA  
GPA: 3.77 / 4.0, Magna Cum Laude

## Honors and Awards

2012  
2017 **National Science Foundation Graduate Research Fellowship**

- Received a three-year fellowship
- Only awardee for Computational Statistics of the thirteen statistical recipients in 2012
- One of five recipients from Iowa State University in 2012

2015/2  
• **Visiting Scientific Researcher**, *Department of Econometrics and Business Statistics*, Monash University - Melbourne, Victoria, Australia

- Received transportation and 1 month of lodging
- Collaborated with [NUMBAT](#) on large data exploration tools

2015/2  
• **Visiting Scientific Researcher**, *Fields Institute for Research in Mathematical Sciences*, Toronto, Ontario, Canada

- Received transportation and lodging
- Spoke at '[Visualization for Big Data: Strategies and Principles](#)' workshop

2011/6  
• **Best Dynamic / Interactive Visualization**, *Data Insight (SF)*

2011/4  
• **National Science Foundation Graduate Research Fellowship**, *Honorable mention*

2008  
2010 **Eta Kappa Nu Honor Society**, *EE/CprE Honor Society for top 25% of class*

2006  
2010 **Dean's List**, *Achieving a 3.5 or above for a semester*

2006  
2010 **Engineering Merit Scholarship**, *Achieving at least a 3.50 GPA in the College of Engineering*

2006  
2010 **President's Award for Competitive Excellence**

## Publications

### Journal Articles

- 2017 Gökalp, F., Barret Schloerke. "Parallel Programming in Linear Mixed Models." *The R Journal*, Submitted 08/2017.
- 2017 Schloerke, B., Hadley Wickham, Dianne Cook, Heike Hofmann. "Escape from Boxland: Generating a Library of High-Dimensional Geometric Shapes." *The R Journal*, 8(2):243-257, December 2016.
- 2013 Emerson, J., W. Green, B. Schloerke, J. Crowley, D. Cook, H. Hofmann, and H. Wickham "The Generalized Pairs Plot." *Journal of Computational and Graphical Statistics* 22.1 (2013). Print.
- 2012 Hofmann H., D. Cook, C. Kielion, B. Schloerke, J. Hobbs, A. Loy, L. Mosley, D. Rockoff, Y. Huang, D. Wrolstad, and T. Yin "Delayed, Canceled, on Time, Boarding... Flying in the USA." *Journal of Computational and Graphical Statistics* 20.2 (2012). Print.
- 2010 Mosley, L., D. Cook, H. Hofmann, C. Kielion, and B. Schloerke. "Monitoring the 2008 Election Visually." *Chance* 23.3 (2010). Print.

### Presentations

- 2016/8 Schloerke, B. "Cognostics: Metrics enabling detailed interactive visualization of big data" CSESC2017. Purdue University, West Lafayette, IN. Aug. 2017. Speech.
- 2017/1 Schloerke, B. "tidyverse[c(\"magrittr\", \"dplyr\", \"tidyr\")]" Purdue Graduate Statistics Seminar. Math, West Lafayette, IN. Jan. 2017. Speech.
- 2016/8 Schloerke, B. "Cognostics: Metrics enabling detailed interactive visualization of big data" Joint Statistical Meeting. Chicago, IL. Aug. 2016. Speech.
- 2016/6 Schloerke, B. "ggduo: Pairs plot for two group data" UseR!2016. Stanford University, Stanford, CA. Jun. 2016. Speech.
- 2016/6 Schloerke, B. "High Performance Computing for High-Resolution Analysis of Big Data and Small" Spring Research Conference. Illinois Institute of Technology, Chicago, IL. Jun. 2016. Speech.
- 2016/4 Schloerke, B. "Analysis and Visualization of Large Complex Data with Tessera". Graduate Statistics Seminar. Purdue, West Lafayette, IN. Mar. 2016. Speech.
- 2016/3 Schloerke, B. "Analysis and Visualization of Large Complex Data with Tessera". Graphics Research Group Seminar. Iowa State University, Ames, IA. Mar. 2016. Speech.
- 2016/2 Schloerke, B. "Analysis and Visualization of Large Complex Data with Tessera". NUMBAT Seminar. Monash University, Melbourne, VIC AUS. Feb. 2016. Speech.
- 2016/2 Schloerke, B. "GraphQL: An R server implementation". Wombat2016. Melbourne Zoo, Melbourne, VIC AUS. Feb. 2016. Speech.
- 2015/10 Schloerke, B. "Web Scraping with R" American Credit Acceptance. Spartanburg, SC. Oct. 2015. Speech.
- 2015/2 Schloerke, B. "Trelliscope: D&R visualization tool" Big Data Visualization. Fields Institute Toronto, CAN. Feb. 2015. Speech.
- 2014/10 Schloerke, B. "ggplot2: displaying spatial and temporal data" Spatial Statistics Seminar. Recitation, West Lafayette, IN. Oct. 2014. Speech.
- 2014/2 Schloerke, B. "Reducing Working Environment Inefficiencies" Graduate Statistics Seminar. Haas, West Lafayette, IN. Feb. 2014. Speech.
- 2010/9 Schloerke, B. "helpR: Help for R." Working Group Meeting. Snedecor Hall, Ames, IA. Sept. 2010. Speech.
- 2010/8 Schloerke, B. "GGally: A Plot Matrix for All Variable Types." Joint Statistical Meeting 2010. Vancouver Convention Center, Vancouver, BC Canada. Aug. 2010. Speech.
- 2009/6 Quart, G., B. Schloerke, and D. Gannon. "Visualizing Large Data Sets: Housing Crisis." VIGRE. Rice University, Houston, TX. June 2009. Speech.

### Talks

- 2014/6 "git", Purdue Working Group

2011  
2011

**“Sinartra: Simple Web Framework”, “5 Tips for Running Node in Production”, “Express: Node Router”, San Francisco Bay Area**

2008  
2011

**“Roxygen: Documentation for R”, “helpr: Help for R”, and “Electoral Towers”, Iowa State University Working Group**

## Research

### Purdue University

Advisors: Dr. Ryan Hafen and Dr. William Cleveland

2015

**trelliscopejs**, *browser data visualization*

- Ported original trelliscope R package to be separated from shiny
- Full visual implementation in javascript dramatically increases interaction time
- Utilized React.js javascript platform to produce a single webpage application with multiple URLs

2015

**GraphQL**, *query language for your API*

- GraphQL is a query language for APIs and a runtime for fulfilling those queries with your existing data. GraphQL provides a complete and understandable description of the data in your API, gives clients the power to ask for exactly what they need and nothing more, makes it easier to evolve APIs over time, and enables powerful developer tools.
- Native R GraphQL implementation to be used with Relay.js as a go-between the browser and a R data backend

2016

**rbokeh**, *R interface to Bokeh*

- A native R plotting library that provides a flexible declarative interface for creating interactive web-based graphics, backed by the Bokeh visualization library.
- Utilized non-standard evaluation within package to remove massive amounts of boilerplate

2016

**GGally::ggduo**, *plot matrix for two grouped data*

- Pairs plot for two group data
- Extension of ggpairs built on top of a ggmatrix object
- Readily useful for multiple time series analysis and canonical correlation analysis
- Foundation function for 'ggnostic', a model diagnostic plot matrix, and 'ggts', a time series plot matrix
- Funded by Google Summer of Code with [final report](#)

2015

**GGally::ggmatrix**, *generalized plot matrix*

- Handles arbitrary plot columns and rows
- Used as display mechanism for all GGally plot matrix functions

2013

**Bitcoin**, *DARPA XDATA*

- Brought previous transaction exporter up to protocol
- Formatted escrow transactions to be matched easier
- Made "None" keys uniquely identifiable, rather than being treated the same
- Matched some "None" keys to valid output keys
- Updated User exporter to include all information available

### Rice University

Advisor: Dr. Hadley Wickham

Summer 2010

**helpr**, *Help for R*

- Better Friendly HTML documentation with links to other packages, function aliases, and function sources. Finding information may be done with the comprehensive search bar.

Summer 2009

**VIGRE Program**

- Performed exploratory analysis on the housing crisis
- Found, cleaned, and explored multiple data sets
- Created all outputs reproducible from scratch

### Novartis Pharmaceutical

Mentors: Dr. Francois Mercier and Dr. Anthony Rossini

Summer 2008

**Novartis Pharmaceutical R&D**

- Created a time management tool using gantt charts to help management quickly display availability.
- Created a web-based tutorial consisting of examples of the most common graphics produced when using different layouts. This allowed users to know exactly how everything was being displayed. (white-box)
- Worked cohesively with the top 15 statisticians within Novartis.

## Iowa State University Statistics Research

Advisors: Dr. Dianne Cook and Dr. Heike Hofmann

2010

### **GGally R package**, *Extension of ggplot2*

- Re-designed the scatterplot matrix within ggplot2 to handle other variable types
- Users may submit custom functions to be displayed in the plot matrix
- Collaborated with over 9 contributing authors
- > 18k downloads a month
- Barret Schloerke, Dr. Wickham, Dr. Cook, Dr. Hofmann, Dr. Hafen
- Iowa State University, Purdue University

2010

2010

### **cranvas**, *New Plotting Canvas for R*

- New plotting canvas for R. Produces more than a million points in less than a second.
- Dr. Wickham, Dr. Lawrence, Dr. Hoffman, Dr. Cook, Yihui Xie, Tengfei Yin, Marie Vendettuoli, Barret Schloerke

2009

2009

### **DescribeDisplay**, *Reproduce GGobi images in R*

- Produce publication quality graphics from GGobi's describe display plugin.

2009

2009

### **tourr**, *Geodesic interpolation and basis generation functions*

- Implements geodesic interpolation and basis generation functions that allow you to create new tour methods in R.
- Dr. Wickham, Dr. Cook, Barret Schloerke
- Iowa State University and Rice University

2009

2009

### **American Statistical Association Data Expo**, *Delayed, Cancelled, On-Time, Boarding ... Flying in the USA*

- Displayed airport delays caused by airport layout using multiple applications.
- Using Google's KML language, R and wget, we produced Google Earth images with the runways highlighted.
- Iowa State University working group finished 2nd in the [world competition](#).

2008

2008

### **Election Data**, *Explore new ways to show the Electoral College for the 2008 Presidential Election*

- Created a custom cartogram of the United States with area proportional to the electoral vote of each state.
- Produced the 'Electoral Vote Towers' (inspired by New York Times 2000 Presidential Race) that quickly portrayed the electoral vote status.

2008

2008

### **DNA**, *Viewing DNA sequence data*

- Extension of the polytope ranking.
- Created combinations of DNA sequences with a length of four which were displayed in a grand tour.

2008

2008

### **Polytope Ranking**, *Visualizing ranked data*

- Created a model that would connect a ranked data set.
- Added jittered data points to each combination with the distance relative to the amount of points.

2007

2007

### **FaceOff**, *'Lines and Dots' in the fourth dimension*

- Game created to calibrate users into viewing and thinking in the fourth dimension.
- Game is a combination of "Connect 4", "Tic Tac Toe", and "Dots and Lines."

2007

2007

### **IEEE Infovis Competition**, *2007 Infovis visualization competition*

- Internet Movie Data Base (IMDB) competition.
- Worked on the genres of the database using the "Many Eyes" web application.
- Helped edit and produce the movie that the working group submitted.

2007

2007

### **GeoZoo**, *A library of high dimension geometric objects*

- Functions to produce cubes, spheres, simplices, polyhedra, polytopes, tori, and mobius-like objects.
- Contains a new addition to the high dimension polytopes, a Hyper Ring Torus

## Work Experience

### Bill and Melinda Gates Foundation Consultant

#### 2016 Tool Development Team

##### Metamarkets - San Francisco, CA

#### 2011-2012 Software Engineer

##### Client-Side

- Developed url framework
- Created the javascript foundation from scratch
- Utilized DVL and D3 to make visualizations and objects data dependant
- Modularized the code to make it useable in multiple locations
- Sped up deployment of code and utilized browser caching
- Minimized calls to backend server to avoid excessive load

##### Server-Side

- Learned and implemented Node.js
- Developed / forked multiple public packages
- Developed user management system
- Developed / utilized user configuration file
- Allowed user configuration to work with self serve process
- Maintained client security
- Implemented custom server cluster to prevent request failure
- Merged multiple data sources per request
- Wrote custom security check and error notification system

#### 2011-2012 DVL, Data Visualization Legos

Functionally reactive javascript library to bind data to functions and visualizations

#### 2012 yalog, Yet Another Logger

Request logger to allow for personalized logging messages within different log levels

#### R Master Class Assistant

Instructor: Hadley Wickham

2015/6 Chicago, IL, "Package Development in R", 50 people

2011/6 San Francisco, CA, "Advanced Programming Techniques" and "Package Development in R", 40 people

## Service

### R for Data Science Seminar

#### 2017 Seminar Instructor, Purdue

- Organize and present course material
- Coordinate with students maximize student attendance
- Field student's questions and adapt presentations to match student abilities

### Graduate Office / Leadership Positions

#### 2016-2017 Graduate Student Office President, Statistics

- Organize Graduate Stent Mentor Program
- Organize Esteemed Speaker event for the Spring semester
- Run elections for the following year

#### 2014-2016 StatCom Assistant

- Coordinated and put on multiple events
- Helped in production of StatCom banner

#### 2014-2016 Graduate Student Mentor

### Undergraduate Office / Leadership Positions

#### 2007-2010 Iowa State University Hip Hop Club (Dub H), Technology Chair and Executive Committee

- Made club self sustaining
- Plan, execute, and review all events
- Produce videos and DVDs for club

#### 2007-2010 GRIP Mentoring, Leader/Friend

Met with and provided guidance for a middle school student in Gilbert, IA

2010  
2010

**Greek Week Check-In, *Electronic Check-in for Events***

- Managed a system with more than 2300 records for 10+ events
- Linked to participants Iowa State University identification card

2009  
2010

**Eta Kappa Nu**

- Organized event to inform middle school students what a Computer Engineer does at Iowa State University  
- Service Chair - Spring 2010
- Oversaw the meetings and made certain that the events were run properly - Vice President - Spring 2009

2007  
2009

**FarmHouse Fraternity, *Scholarship Chair***

- Implemented five new programs to strengthen GPA
- Moved committee chair into an executive position

2007  
2007

**E-Ball Subcommittee, *Committee Member***

- Planned dance event for 800 people
- Specialized in food selection

---

## References

Dr. William Cleveland

Professor

Statistics, Purdue University

✉ [wsc@purdue.edu](mailto:wsc@purdue.edu)

Dr. Ryan Hafen

Hafen Consulting

Adjunct, Purdue University

✉ [rhafen@gmail.com](mailto:rhafen@gmail.com)

Dr. Hadley Wickham

Data Scientist

RStudio

✉ [h.wickham@gmail.com](mailto:h.wickham@gmail.com)

Dr. Dianne Cook

Professor

Monash Univeristy

✉ [dicook@monash.edu](mailto:dicook@monash.edu)

Dr. Heike Hofmann

Professor

Iowa State University

✉ [hofmann@iastate.edu](mailto:hofmann@iastate.edu)

Dr. Michael Driscoll

CEO

Metamarkets

✉ [mike@metamarkets.com](mailto:mike@metamarkets.com)