Barret Schloerke

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

✉ schloerke@gmail.com
in schloerke
 schloerke

Education

2013 – 2017

Doctor of Philosophy in Statistics, *Purdue University*, West Lafayette, IN
Advisors: Dr. William Cleveland and Dr. Ryan Hafen
National Science Foundation Graduate Research Fellowship Recipient (2012-2017)

2013 – 2014

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

2012 – 2012

Doctor of Philosophy in Statistics, *Rice University*, Houston, TX
Advisor: Dr. Hadley Wickham

2006 – 2010

Bachelor of Science in Computer Engineering, *Iowa State University*, Ames, IA, 3.77/4.0

Work Experience

2011 – 2012

Metamarkets, *San Francisco, CA*, Software Engineer

Client-Side (`Coffeescript`)	Front-end Server (`Node.js & Coffeescript`)
○ Implemented URL routing framework which displayed consistent URLs between users	○ Automated server deployment process to allow any engineer to deploy services
○ Implemented DVL visualizations using D3 for all interactive data visualizations	○ Created server cluster to prevent request failure
○ Utilized DVL library to make all variables reactive to dependency value changes	○ Developed company/user management system
○ Modularized code to be reusable / plug-in-play	○ Developed configuration files to determine where and how data is displayed
○ Reduced website startup time for faster loading	○ Integrated configuration files to help automate self-serve data processing
○ Trimmed data queries to grab new data only to reduce unnecessary server load	○ Maintained client security
○ Refined anomaly detection algorithms to highlight visual anomalies	○ Communicated with multiple back-end services to maintain a consistent client interface

2016

Bill and Melinda Gates Foundation Consultant, *Remote*, Tool Development Team
○ Provided customized summary statistics in timely manner for 75+ datasets to feed into visualization applications
○ Maintained data security while coordinating with 5 other data scientists

2016 – 2017

DARPA: XDATA and D3M Grants, *Purdue University*
○ Collaborated remotely with Hafen Consulting, Kitware, and KnowledgeVis
○ Completed multiple miniature hack-a-thons with my team to present a consistent findings of our results
○ Served as a point of contact to integrate services with other working groups

2017 – 2017

Big Data Analytics: Statistics and Data Visualization, *Purdue University*, Teaching Assistant
○ Taught several lectures to help students understand how to use R for *big* data
○ Explored easy to understand concepts in class that have difficulty scaling to larger data

2011 2015

Hadley Wickham R Master Class, *San Francisco, CA and Chicago, IL*, Teaching Assistant
○ Fielded advanced programming technique questions from students to solidify content presented
○ Aided students in developing R packages on multiple computing platforms

2008

Novartis Pharmaceutical R&D, *Basel, Switzerland*, Data Scientist Intern
○ Created a time management tool using gantt charts to help management quickly display employee availability
○ Created a web-based tutorial consisting of examples of the most common graphics produced when using different layouts by "white-box"ing complex visualizations into reusable plotting layers
○ Worked cohesively with the top 15 statisticians within Novartis

Honors and Awards

National Science Foundation Graduate Research Fellowship
- Received a three-year fellowship (to be used within five years) to pursue graduate studies in a science based field
- Only awardee for Computational Statistics of the thirteen statistical recipients in 2012
- One of five recipients from Iowa State University in 2012

Visiting Scientific Researcher, *Monash University*, Melbourne, Victoria, Australia
- Presented latest research on `gqlr` to WOMBAT2016
- Collaborated with NUMBAT on large data exploration tools

Best Dynamic / Interactive Visualization, *Data Insight (SF)*, San Francisco, CA
- Team of four worked together over the weekend to produce an interactive visualization
- Leaned on each other's skills to provide an application that highlighted our data and visualization talents

National Science Foundation Graduate Research Fellowship, *Honorable mention*

Publications

Journal Articles

Gökalp, F., Barret Schloerke. "Parallel Programming in Linear Mixed Models." *The R Journal*, Submitted 08/2017.

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.

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.

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.

Mosley, L., D. Cook, H. Hofmann, C. Kielion, and B. Schloerke. "Monitoring the 2008 Election Visually." *Chance* 23.3 (2010). Print.

Presentations

Schloerke, B. "Cognostics: Metrics enabling detailed interactive visualization of big data" CSESC2017. Purdue University, West Lafayette, IN. Apr. 2017. Speech.

Schloerke, B. "tidyverse[c("magrittr", "dplyr", "tidyr")]" Purdue Graduate Statistics Seminar. Math, West Lafayette, IN. Jan. 2017. Speech.

Schloerke, B. "Cognostics: Metrics enabling detailed interactive visualization of big data" Joint Statistical Meeting. Chicago, IL. Aug. 2016. Speech.

Schloerke, B. "ggduo: Pairs plot for two group data" UseR!2016. Stanford University, Stanford, CA. Jun. 2016. Speech.

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.

Schloerke, B. "Analysis and Visualization of Large Complex Data with Tessera". Graduate Statistics Seminar. Purdue, West Lafayette, IN. Mar. 2016. Speech.

Schloerke, B. "Analysis and Visualization of Large Complex Data with Tessera". Graphics Research Group Seminar. Iowa State University, Ames, IA. Mar. 2016. Speech.

Schloerke, B. "Analysis and Visualization of Large Complex Data with Tessera". NUMBAT Seminar. Monash University, Melbourne, VIC AUS. Feb. 2016. Speech.

Schloerke, B. "GraphQL: An R server implementation". Wombat2016. Melbourne Zoo, Melbourne, VIC AUS. Feb. 2016. Speech.

Schloerke, B. "Web Scraping with R" American Credit Acceptance. Spartanburg, SC. Oct. 2015. Speech.

Schloerke, B. "Trelliscope: D&R visualization tool" Big Data Visualization. Fields Institute Toronto, CAN. Feb. 2015. Speech.

Schloerke, B. "ggplot2: displaying spatial and temporal data" Spatial Statistics Seminar. Recitation, West Lafayette, IN. Oct. 2014. Speech.

Schloerke, B. "Reducing Working Environment Inefficiencies" Graduate Statistics Seminar. Haas, West Lafayette, IN. Feb. 2014. Speech.

Schloerke, B. "helpr: Help for R." Working Group Meeting. Snedecor Hall, Ames, IA. Sept. 2010. Speech.

Schloerke, B. "GGally: A Plot Matrix for All Variable Types." Joint Statistical Meeting 2010. Vancouver Convention Center, Vancouver, BC Canada. Aug. 2010. Speech.

Lightning Talks

"git", Purdue Working Group

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

Research

Purdue University

Advisors: Dr. Ryan Hafen and Dr. William Cleveland

`autocogs`, Automatic Cognostic Calculations, *Purdue University*
- "Automatically calculates cognostics for plot objects and list column plot objects. `autocogs` compliments `trelliscopejs`'s panel interactions by producing multiple cognositc values for the visualizations displayed"
- Generalized framework to produce consistent cognostics independent of visualization library utilized

`gqlr`, GraphQL Server in R, *Purdue University*
- "R server implementation of GraphQL, a query language created by Facebook for describing complex data queries independent of the storage format."
- GraphQL provides a complete and human readable description of the data in your data schema and gives clients the power to query only for exactly what they need.
- gqlr is a native R GraphQL implementation to be used with `Relay` data management in a `React` javascript web application

`GGally::ggduo`, Generalized plot matrix for two-grouped data, *Purdue University*
- Extension of ggpairs built on top of a ggmatrix object
- Pairs plot for two group data
- 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

`GGally::ggmatrix`, Generalized plot matrix, *Purdue University*
- Handles arbitrary plot columns and rows
- Used as display mechanism for all GGally plot matrix functions

`trelliscopejs`, Create Interactive Trelliscope Displays, *Purdue University*
- "Trelliscope is a scalable, flexible, interactive approach to visualizing data. This package provides methods that make it easy to create a Trelliscope display specification for `trelliscopejs`. High-level functions are provided for creating displays from within `dplyr` or `ggplot2` workflows. Low-level functions are also provided for creating new interfaces."
- Ported Shiny-based `trelliscope` R package to be built with React framework to increase interaction speed
- Integrated with `ggplot2` objects to seamlessly produce `trelliscopejs` applications

`rbokeh`, R interface to Bokeh, *Purdue University*
- "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

Bitcoin, *DARPA XDATA*
2013
2013

- 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

Metamarkets

DVL, Data Visualization Legos, *Metamarkets*
2011
2012

- Functionally reactive javascript library to bind data to functions and visualizations

yalog, Yet Another Logger, *Metamarkets*
2012
2012

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

Rice University

Advisor: Dr. Hadley Wickham

helpr, *Help for R*
2010
2010

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

VIGRE Program
2010
2010

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

Iowa State University Statistics Research

Advisors: Dr. Dianne Cook and Dr. Heike Hofmann
2010

GGally, *Extension to ggplot2, Iowa State University, Purdue University, and Google Summer of Code*

- "GGally extends ggplot2 by adding several functions to reduce the complexity of combining geometric objects with transformed data. Some of these functions include a pairwise plot matrix, a two group pairwise plot matrix, a parallel coordinates plot, a survival plot, and several functions to plot networks."
- Maintained R package with 25k+ monthly downloads
- Collaborated with 10 major authors and many contributors
- Created multiple plot matrix functions to aid in full data exploration
- Integrated development with lintr and testthat for code and output consistency

cranvas, *New Plotting Canvas for R*
2010
2010

- 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

DescribeDisplay, *Reproduce GGobi images in R*
2009
2009

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

tourr, *Geodesic interpolation and basis generation functions*
2009
2009

- 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

American Statistical Association Data Expo, *Delayed, Cancelled, On-Time, Boarding ... Flying in the USA*
2009
2009

- 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.

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

- 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, polytops, tori, and mobius-like objects.
- Contains a new addition to the high dimension polytopes, a Hyper Ring Torus

Service

2017
2017

R for Data Science Seminar, *Purdue University*, Instructor
- Organized and presented course material in interactive weekly sessions
- Answered student's questions and adapted the presentations accommodate variable skill levels
- Coordinated with students maximize student attendance

2016
2017

Statistics Graduate Student Office, *Purdue University*, President
- Effectively ran a town hall meeting to positively address major concerns of statistics graduate students
- Organized "Graduate Student Mentor Program" at beginning of Fall Semester
- Organized "Esteemed Speaker" event with Dr. Arthur Dempster for the Spring semester

2014
2016

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

2014
2016

Graduate Student Mentor

Extra

Languages
- Expert: R, Javascript (ES5), Node.js, GraphQL, Markdown, HTML, LaTeX, JSON, YAML
- Moderate: MySQL, JSX, Regular Expressions, Bash, CSS, C

Systems
- Expert: Travis CI, GitHub, GitHub Pages, `tidyverse.R`, Hadoop
- Moderate: `React.js`, `Drat.R`, `Broccoli.js`, `Webpack.js`