

Employment

- 2013– Chief Scientist, RStudio.
- 2013– Adjunct Professor, Rice University, Houston, TX.
- 2008–12 Assistant Professor, Rice University, Houston, TX.

Education

- 2008 Ph.D. (Statistics), Iowa State University, Ames, IA. “Practical tools for exploring data and models.”
- 2004 M.Sc. (Statistics), First Class Honours, The University of Auckland, Auckland, New Zealand.
- 2002 B.Sc. (Statistics & Computer Science), First Class Honours, The University of Auckland, Auckland, New Zealand.
- 1999 Bachelor of Human Biology, First Class Honours, The University of Auckland, Auckland, New Zealand.

Honours & awards

- 2011 Distinguished faculty associate, McMurtry College.
- 2010 Best paper award Infovis '10 for “Graphical inference for infovis”. With Dianne Cook, Heike Hofmann and Andreas Buja.
- 2010 Distinguished faculty associate, McMurtry College.
- 2007 Student paper award for “Meifly: model exploration with GGobi and R”. ASA Computing and Graphics Sections.
- 2007 Vince Sposito award for outstanding potential in statistical computing. Department of Statistics, Iowa State University.
- 2007 Preparing Future Faculty Scholar, Iowa State University.
- 2006 John M. Chambers Statistical Software Award for “ggplot and reshape: Practical tools for organising, summarising and exploring data”. ASA Computing and Graphics Sections.
- 2006 Second place. ASA Data Expo Challenge, ASA Computing and Graphics Sections. With Jon Hobbs, Heike Hofmann and Dianne Cook.
- 2006 Dan Mowrey Statistical Consulting for excellence in consulting. Department of Statistics, Iowa State University.
- 2005 First place. Infovis 2005 data visualisation contest. With Heike Hofmann, Dianne Cook, Christian Röttger, and Junjie Sun.
- 2004–8 Miller Fellowship. Department of Statistics, Iowa State University.

Publications

Books

- 2015 H. Wickham. “R Packages” (In preparation). To be published by O’Reilly.
- 2014 H. Wickham. “Advanced R” (2014). Published by Chapman & Hall.

2009 H. Wickham. “ggplot2: Elegant Graphics for Data Analysis” (2009). Published by Springer.

2007 Contributor to Cook & Swayne “Interactive and Dynamic Graphics for Data Analysis: With Examples Using R and GGobi” (2007).

Journal articles

TBA G. Grolemond and H. Wickham (Accepted). “Visualizing Complex Data with Embedded Plots”. In: *Journal of Computational and Graphical Statistics*

2014 **tidy-data** [Downloaded over 10,000 times]

2014 G. Grolemond and H. Wickham (2014). “A Cognitive Interpretation of Data Analysis”. In: *International Journal of Statistics* 82.2, pp. 184–204 [Discussants: Leland Wilkinson & Peter Huber].

2013 G. J. Spiegel et al. In:

2013 **stat-graph-hist**

2013 D. Kahle and H. Wickham (2013). “ggmap: Spatial Visualization with ggplot2”. In: *The R Journal* 5 (1), pp. 144–162. URL: <http://journal.r-project.org/archive/2013-1/kahle-wickham.pdf>.

2012 **glyph-maps**

2012 J. W. Emerson et al. (2012). “The Generalized Pairs Plot”. In: *Journal of Computational and Graphical Statistics* 22.1, pp. 79–91. URL: <http://www.tandfonline.com/doi/ref/10.1080/10618600.2012.694762>.

2012 B. Huang, D. Cook, and H. Wickham (To appear). “A Tour GUI using gWidgets”. In: *JSS*.

2011 **mutatr**

2011 **me:prodplots**

2011 G. Quart and H. Wickham (2011). “Protection Against the Housing Crisis: An Analysis of Industrial Diversity”. In: *The Rice Cultivator: A Student Journal on Public Policy* 2, pp. 69–77.

2011 **me:testthat**

2011 **me:data-expo-09**

2011 **me:ggplot2-wires**

2011 **me:plyr** [Downloaded over 99,000 times.]

2011 **me:tourr** [Downloaded over 8,000 times.]

2011 G. Grolemond and H. Wickham (2011). “Dates and Times Made Easy with lubridate”. In: *Journal of Statistical Software* 40.3, pp. 1–25. URL: <http://www.jstatsoft.org/v40/i03/>. [Downloaded over 37,000 times.]

2010 **me:inf4info**

2010 **me:stringr**

2010 D. M. Debinski et al. (2010). “Montane meadow change during drought varies with background hydrologic regime and plant functional group”. In: *Ecology* 91.6, pp. 1672–1681.

2009 **wickham:2007d**

2009 J. Hobbs et al. (2010). “Glaciers Melt as Mountains Warm: A Graphical Case Study”. In: *Computational Statistics* 25.4. Special issue for ASA Statistical Computing and Graphics Data Expo 2007. In press, pp. 569–586.

2009 **wickham:2008a**

2009 M. Lawrence et al. (2009). “Extending the GGobi pipeline from R: Rapid Prototyping of Interactive Visualizations”. In: *Computational Statistics* 24.2. Special issue for Proceedings of the 5th International Workshop on Directions in

Statistical Computing., pp. 195–205.

2009 A. Buja et al. (2009). “Statistical Inference for Exploratory Data Analysis and Model Diagnostics”. In: *Royal Society Philosophical Transactions A* 367.1906, pp. 4361–4383.

2008 D. Cook and H. Wickham (2008). “Comment on The Future of Statistical Computing”. In: *Technometrics* 50.4, pp. 442–443.

2007 **wickham:2007b** [Downloaded over 32,000 times.]

2006 A. Oakley et al. (2006). “Diagnostic value of written referral and/or images for skin lesions”. In: *Journal of Telemedicine and Telecare* 12, pp. 151–158.

2000 **wickham:2000**

Book chapters & encyclopedia articles

2012 **r-s**

2012 **stat-graphics**

2009 **wickham:2009**

2008 D. Caragea et al. (2008). “Visual Methods for Examining SVM Classifiers”. In: *Visual Data Mining*. Ed. by S. Simoff, M. H. Böhlen, and A. Mazeika. LNCS State of the Art Surveys, pp. 136–153.

2006 D. Cook et al. (2008). “Grand Tours, Projection Pursuit Guided Tours and Manual Controls”. In: *Handbook of Data Visualization*. Ed. by C. Chen, W. Härdle, and A. Unwin. Springer Handbooks of Computational Statistics. Springer. Chap. III.2, pp. 295–314.

Posters

2011 H. Deng and H. Wickham. “Density estimation in R” *Rice Undergraduate Research Symposium*, April 2011.

2011 H. Wickham and G. Grolemond. “Teaching data analysis with statistical thinking and visualization”, *Scientia Conference on Research and Innovation in Undergraduate Science and Engineering Education*, February 2011.

2010 G. Quart and H. Wickham. “Data visualisation - industrial protection against the housing crisis” *Rice Undergraduate Research Symposium*, April 2010.

2009 G. Quart, B. Schloerke, D. Gannon, G. Grolemond and H. Wickham. “Visualization of large data sets - the housing crisis”, *VIGRE 2009 Info/Poster Session*, October 2009.

2007 H. Hofmann, D. Cook, U. Genschel, H. Wickham, M Lawrence, B Schloerke, and S. Bradley, “Bring your popcorn and enjoy the show!” *Infovis*, 2007. <http://had.co.nz/infovis-2007/>.

2006 J. Hobbs et al. (2006). “Another View at Central America”. In: *JSM*. <http://had.co.nz/dataexpo>. URL: <http://had.co.nz/dataexpo>.

2006 **wickham:2006f**

Book reviews

2010 **me:review-gsdar**

2010 **me:review-msod**

2008 **me:review-stat-prog**

2008 **me:o8-data-manip**

2007 **wickham:2007aa**

Editorials

- 2010 R. A. Levine et al. (2010). “Editorial: Publishing Animations, 3D Visualizations, and Movies in JCGS”. in: *Journal of Computational and Graphical Statistics* 19.1, pp. 1–2. eprint: <http://pubs.amstat.org/doi/pdf/10.1198/jcgs.2010.191ed>.

Unrefereed publications

- 2013 J. Huchette and H. Wickham (2013). *A constraint-based layout approach to data visualization*. Tech. rep. had.co.nz.

2013 **bigvis**

2013 **boxplots**

- 2011 H. Deng and H. Wickham (2011). *Density estimation in R*. tech. rep. had.co.nz

2008 **wickham:2008b**

2008 **wickham:2008b**

2007 **weisburg:2007**

2007 **wickham:2007h**

2007 **me:dsc**

- 2006 H. Hofmann, K. Kafadar, and H. Wickham (2006). *Letter-Value Box Plots - adjusting box plots for large data sets*. Tech. rep. 10. Department of Statistics, Iowa State University. URL: <http://www.stat.iastate.edu/preprint/articles/2006-10.pdf>.

2006 **wickham:2006**

2005 **wickham:2005a**

Talks (2013–)

Meetups & user groups

- 2014 Sep 2014, Data Visualisation NYC. “Introducing ggvis”.
Mar 2014, Seattle. “ggvis sneak peek”.
Mar 2014, Portland. “ggvis sneak peek”.
2013 Oct 2013, Seattle. “Introducing dplyr”.
Sep 2013, Chicago. “Introducing dplyr”.
Jul 2013. London. “BiggeR data analysis”, <http://bit.ly/bigrdata3>.
Jun 2013. Dublin. “BiggeR data analysis”, <http://bit.ly/bigrdata>.
Jun 2013. Twin cities. “R packages are easy”, <http://bit.ly/pkgsrez3>.
Jun 2013. Bay Area. “R packages are easy”, <http://bit.ly/pkgsrez>.
May 2013. Vancouver. “R packages are easy”, <http://bit.ly/pkgsrez>.
May 2013, New York. “Visualising 100,000,000 observations with R”.

Conferences

- 2014 Sep 2014, Effective Applications of the R language. “Expressing yourself in R” [keynote].

Sep 2014, Strata. “Expressing yourself in R”.

Aug 2014, JSM. “ggvis: Moving Toward a Grammar of Interactive Graphics”.

Jul 2014, Australian Statistics Conference. “Visualising bigger data in R”.

Jul 2014, UseR! “Data manipulation with dplyr”.

Jun 2014, Computational Methods for Massive/Complex Data. “Data tidying and munging”.
<http://www3.imperial.ac.uk/statistics/complexdata/videos>

May 2014, Chinese R User’s Conference, “R packages: principles and best practices” [keynote].

May 2014, Chinese R User’s Conference, “ggvis sneak peek”.

Apr 2014, Paypal SkySong. “Expressing yourself in code”.

Feb 2014, Strata. “Expressing yourself in R”.

Jan 2014, Data Centric Programming. “A grammar for data analysis”.

2013 Nov 2013, NSF DALI. “Data analysis bottlenecks”.

Oct 2013, Simply Statistics Unconference. “The future of statistical software”, <http://bit.ly/futurestatsoft>.

Aug 2013, JSM. “BiggerR data analysis”. <http://bit.ly/bigdata4>.

Jul 2013, UseR 2013! “BiggerR data analysis” [keynote]. <http://bit.ly/bigdata2>.

Jul 2013, Rencontres R. “Visualising big data (in R)” [keynote]. <http://bit.ly/bigdata1>.

Mar 2013, Conference of Texas Statisticians.

Mar 2013, Biometrics.

Mar 2013, Investigative reporting and editing conference. “Data science for the perplexed”.

Academic visits

2014 Mar 2014, Oregon State University.

Apr 2014, Harvard.

Apr 2014, Google Boston.

Feb 2014, Stanford. “Expressing yourself in R”. <https://www.youtube.com/watch?v=wki0BqlztCo>.

2013 Oct 2013, Google. “Expressive data analysis”.

Sep 2013, MIT. “Visualising 100,000,000 and beyond: design principles”.

Sep 2013, Purdue (Statistics). “Data analysis bottlenecks”.

Sep 2013, National Opinion Research Center. “Visualising 100,000,000 and beyond: design principles”.

Jul 2013, University City London. “BiggerR data analysis”.

May 2013, UBC. “Visualising big data in R”.

May 2013, US Census.

Apr 2013, UCLA (Political Science, Finance)

Other

2013 Oct 2013. Scientia, Rice. “Investing in your intellect”, <http://bit.ly/1hecEX7>.

Talks (2006–2012)

Keynotes

- 2011 “Tidy data”. Commonwealth Scientific and Industrial Research Organisation (CSIRO) R tech fest, Canberra, Australia. October 2011.

Invited

- 2012 “Perspectives on Visualisation.” Scientific Computing Institute, University of Utah, Salt Lake City, UT. April 2012.
- 2011 “Creating effective graphics.” AT&T Research, Summit, NJ. December 2011.
- 2011 “Tidy data”. New York Data Hackers Group, New York City, NY. December 2011. [Over 200 signed up to attend.]
- 2011 “Creating effective graphics”. Bit.ly tech talk, New York City, NY. December 2011.
- 2011 “Creating effective graphics”. News Corp, New York City, NY. December 2011.
- 2011 “Tidy data”. Statistics Department, CMU, Pittsburg, PA. December 2011.
- 2011 “Tidy data”. Department of Statistics, University of Auckland, Auckland, New Zealand. November 2011.
- 2011 “Tidy data”. Sydney R users group, Sydney, Australia. October 2011.
- 2011 “Model visualisation: removing the blindfold”. CMIS, CSIRO, Sydney, Australia. October 2011.
- 2011 “Using ensembles of simple models to understand the Texas housing market”. Eubanks Conference, Rice University, Houston, Texas. September 2011.
- 2011 “tourr: An R package for exploring multivariate data with projections”. ISI, Dublin, Ireland. August 2011.
- 2011 “What’s next for interactive graphics in R”. Statistics 2011, Montreal, Canada. July 2011.
- 2011 “Big data and R”. Ebay, Palo Alto, CA. June 2011.
- 2011 “Engineering data analysis”. Facebook, Palo Alto, CA. June 2011.
- 2011 “What’s next for interactive graphics in R”. Joint meeting for the Bay Area R users and visualisation groups. Mountain View, CA. June 2011. [Over 200 signed up to attend.]
- 2011 “Engineering data analysis”. Google tech talk, Mountain View, CA. June 2011.
- 2011 “Graphical inference”. Lehman Symposium, Rice University, Houston, TX. May, 2011.
- 2010 “Model visualisation: removing the blindfold”. Columbia University, New York, NY. September 2010.
- 2010 “Balance”. Part of “Negotiating the ideal faculty position” workshop, Rice University. With J. Suh, C. Masiello and J. Silberg. September 2010.
- 2010 “ggplot2: a layered grammar of graphics”. *Joint Statistical Meetings*, Vancouver, Canada. July 2010.
- 2010 “An Interactive Graphics Framework for R”. Presented by Michael Lawrence. *Joint Statistical Meetings*, Vancouver, Canada. July 2010.
- 2010 “Teaching data analysis”. *International Conference on Teaching Statistics*, Ljubljana, Slovenia. June 2010.
- 2010 “Managing your email”. Postdoctoral lunch series, Rice University. May 2010.
- 2010 “5 W’s and an H of visualising data”. Ken Kennedy Institute luncheon talk, Rice University, Houston, TX. February 2010
- 2009 “Graphical critique & theory”. Southern California Chapter of the American Statistical Association Fall Kick-Off, Los Angeles, CA. November 2009.

- 2009 “Exploring the housing crisis with ggplot2 and plyr”. Southern California Chapter of the American Statistical Association Fall Kick-Off, Los Angeles, CA. November 2009.
- 2009 “Exploring the housing crisis with ggplot2 and plyr”. RAND Corporation, Santa Monica, CA. November 2009.
- 2009 “The academic vita”, with Junghae Suh. Postdoctoral lunch series, Rice University. October 2009.
- 2009 “Exploring the housing crisis with ggplot2 and plyr”. Bay Area R users group, Berkeley, CA. September 2009.
- 2009 “Butterflies as indicators of environmental change in montane meadows”. Debinski, D. M., J. Caruthers, and H. Wickham. VIth International Conference on Arthropods: Chemical, Physiological, Biotechnological and Environmental Aspects, Ochotnica Dolna, Poland. June 2009.
- 2009 “A graphical grammar + graphical inference = a grammar of graphical inference?” ISU 75th Statistical Laboratory Conference, Iowa State University, Ames, IA. June 2009.
- 2009 “A grammar of graphics; past, present and future”. University of Waikato, Hamilton, New Zealand. May 2009.
- 2009 “Plyr, one data analytic strategy”. University of Auckland, Auckland, New Zealand. May 2009.
- 2009 “Managing your email”. Postdoctoral lunch series, Rice University. May 2009.
- 2009 “The academic vita”, with Junghae Suh. Postdoctoral lunch series, Rice University. February 2009.
- 2008 “Visualising statistical models: Removing the blindfold”. University of Houston, Houston, TX. December 2008.
- 2008 “Visualising statistical models: Removing the blindfold”. Houston Area Chapter of the American Statistical Association, Houston, TX. September 2008.
- 2008 “An introduction to (the new) CIS,” with J. Pitman. Introductory Overview Lecture, *Joint Statistical Meetings*, Denver, CO. August 2008.
- 2008 “Efficiently storing and reshaping large data”. Visualising large data session. *Joint Statistical Meetings*, Denver, CO. August 2008.
- 2008 “Inference for data graphics”. *World Congress in Probability and Statistics*, Singapore. July 2008.
- 2007 Preparing Future Faculty panel. Centre for Excellence in Learning and Teaching, Iowa State University, Ames, IA. November 2007.
- 2007 “Removing the blindfold: visualising statistical algorithms with R & GGobi.” UC Berkeley, Berkeley, CA. October 2007.
- 2007 “scagnostics,” with D. Temple Lang. *Joint Statistical Meetings*, Salt Lake City, UT. August 2007.
- 2007 “meifly: exploratory model analysis.” *Joint Statistical Meetings*, Salt Lake City, UT. August 2007.
- 2007 “ggplot: new software, old ideas.” *useR!*, Ames, IA. August 2007.
- 2007 “An introduction to ggplot.” Novartis AG, Basel, Switzerland. July 2007.
- 2007 “I sold my family secrets and all I got was \$15.” *IgniteIt!*, part of Emerging Technologies Conference, Ames, IA. Video available at <http://vimeo.com/202394>. April 2007.
- 2007 “Interactive and dynamic graphics for animal breeding.” *Animal Breeding and Genetics Seminar*, Department of Animal Science, Iowa State University, Ames, IA. January 2007.
- 2006 “Principles and paradigms of interactive graphics.” The University of Waikato, Hamilton, New Zealand. November 2006.
- 2006 “A Grammar of Graphics.” The University of Auckland, Auckland, New Zealand. November 2006.
- 2006 “A Grammar of Graphics.” AT&T Research Labs, Summit, NJ. October 2006.
- 2006 “rggobi.” UC Berkeley, Berkeley, CA. May 2006.

- 2006 “ggplot, an implementation of the grammar of graphics in R.” University of Augsburg, Augsburg, Germany. May 2006.
- 2005 “Boom and Bust of High-Tech Industry at the turn of the Millenium,” with H. Hofmann. *InfoVis*, Minneapolis, MN. October 2005.

Panels

- 2012 “Plus ça change, plus ça meme chose”. Data science summit, Las Vegas, NV. May 2012.
- 2011 “Teaching data analysis”, JSM, Miami, FL. August 2011.

Contributed

- 2010 “GGally: A plot matrix for all variable types”. Barret E. Schloerke, Dianne Cook, **Hadley Wickham**, Heike Hofmann. *Joint Statistical Meetings*, Vancouver, Canada. July 2010.
- 2010 “qtpaintgui(): rapidly generating interactive graphics in R”. Marie Vendettuoli, Michael Lawrence, **Hadley Wickham**, Dianne Cook, Heike Hofmann. *Joint Statistical Meetings*, Vancouver, Canada. July 2010.
- 2009 “Yet another object oriented system for R.” *Directions in Statistical Computing*, Copenhagen, Denmark. July 2009.
- 2009 “Model visualisation with ggplot2.” *useR! 2009*, Rennes, France. July 2009.
- 2008 “Quantifying ecological effects of changing precipitation on plant communities in montane meadows.” Debinski, D.M., Wickham, H. and J. Caruthers. International Association of Landscape Ecologists. Madison, WI, April 6-10, 2008.
- 2007 “ggplot2: an implementation of the grammar of graphics in R.” *Statistical Computing*, Reimensburg, Günzburg, Germany. July 2007.
- 2006 “Interactive and dynamic graphics for animal breeding,” *World Congress on Genetics Applied to Livestock Production*, Belo Horizonte, Brazil. August 2006.
- 2006 “An implementation of the grammar of graphics in R: ggplot,” *Joint Statistical Meetings*, Seattle, WA. August 2006.
- 2006 “Letter value box plots: Box plots for large datasets,” with K. Kafadar, H. Hofmann. *Joint Statistical Meetings*, Seattle, WA. August 2006.
- 2006 “Data checking with reshape and ggplot.” *Statistical Computing*, Reimensburg, Günzburg, Germany. July 2006.
- 2006 “ggplot software demonstration.” *Statistical Computing*, Reimensburg, Günzburg, Germany. July 2006.
- 2006 “The letter value plot,” with H. Hofmann, K. Kafadar. *UseR!*, Vienna, Austria. June 2006.
- 2006 “The second date: GGobi + R,” with M. Lawrence. *UseR!*, Vienna, Austria. June 2006.
- 2006 “ggplot: An implementation of the grammar of graphics in R.” *UseR!*, Vienna, Austria. June 2006.
- 2006 “Exploring high-dimensional classification boundaries,” with D. Caragea, and D. Cook. *Interface*, Pasadena, CA. May 2006.
- 2006 “Exploratory functional data analysis.” *VIGRE Graphical and Computational Statistics Group*, Department of Statistics, Iowa State University. March 2006.
- 2006 “ggplots: a new graphics package for R” to *VIGRE Graphical and Computational Statistics Group*, Department of Statistics, Iowa State University. February 2006.
- 2004 “marrayVis: Normalisation and Visualisation of Microarrays” to *VIGRE Graphical and Computational Statistics Group*, Department of Statistics, Iowa State University. September 2004.
- 2004 “marrayVis: Normalisation and Visualisation of Microarrays” to *VIGRE Bioinformatics Group*, Department of Statistics, Iowa State University. September 2004.

Invited conference & workshops

- 2012 “Foo camp.” Sebastopol, CA. June 2012.
- 2012 Schloss Dagstuhl, Leibniz Center for Informatics. Dagstuhl, Germany. February 2012.
- 2012 Sparkcamp. Austin, TX. January 2012.
- 2011 “Foo camp.” Sebastopol, CA. June 2011.
- 2009 “Visualising the past.” Richmond, VA. February 2009.
- 2007 “Visualization,” Stephen Few. Kandersteg, Switzerland. July 2007.
- 2004 “MSRI Hot topics: Mathematical and Statistical Methods for Visualization and Analysis of High Dimensional Data.” Berkeley, CA. December 2004.

Software

ggplot2

- 2006– Author of ggplot2, <http://had.co.nz/ggplot2>, an R package for producing statistical graphics. It is unlike most other graphics packages because it has a deep underlying grammar. This grammar, based on the Grammar of Graphics, is made up of a set of independent components that can be composed in many different ways. This makes ggplot2 very powerful, because you are not limited to a set of pre-specified graphics, but you can create new graphics that are precisely tailored for your problem.

The ggplot2 book has been translated into Japanese (2011) and Chinese (2012).

GGobi

- 2006–10 Contributor to the GGobi project, <http://www.ggobi.org>, along with Deborah F Swayne, Dianne Cook, Duncan Temple Lang, Andreas Buja, Heike Hofmann, and Michael Lawrence. My work focused on improving the infrastructure of GGobi to make it easier to add new features in future: more basis generation functions for the tour, an improved projection pursuit algorithm, and randomisation tools to support graphical inference. I also worked on improving the scripting capabilities of GGobi, so that it can easily be controlled from other programming languages.

Other R packages

- 2014– dplyr, an update to plyr with pluggable backends (data frames, data tables, databases, ...), massively improved performance for in-memory data, and an improved grammar.
- 2014– tidyr, an update to reshape2 with improved grammar.
- 2014– babynames, fueleconomy, nasaweather and nycflights13, data packages containing large, interesting, interlinked data frames.
- 2014– rvest, easily scrape data off webpages.
- 2013– pryr, to pry back the covers of R and explore the depths of the language. Available from <https://github.com/hadley/pryr>.
- 2012– staticdocs, an easy way to publish R package documentation online. Available from <https://github.com/hadley/staticdocs>.
- 2012– httr, making http requests easy in R. Available from <https://github.com/hadley/httr>.
- 2011– roxygen2, a tool for easier documentation in R. Available from CRAN.
- 2011– scales, scale functions for visualisation. Available from CRAN.

2010 decumar: an alternative to sweave for interleaving latex and R results. Available from <http://github.com/hadley/decumar>.

2010- devtools: make it easier to develop packages in R. Available from CRAN.

2010 evaluate: parsing and evaluation tools that provide more details than the default. Available from CRAN.

2010 helprr: an alternative documentation interface. With Barret Schloerke. Available from <http://github.com/hadley/helprr>.

2010 lubridate: make working with dates easy and fun. With Garrett Grolemund. Available from CRAN.

2010 memoise: simple memoisation. Available from CRAN.

2010 nullabor: create and display null datasets for graphical inference. Available from <http://github.com/hadley/nullabor>.

2010 Contributor to tourrGui, a graphical user interface to control the tour. Available from CRAN.

2009- testthat: a testing package specifically tailored for R that's fun, flexible and easy to set up. Available from CRAN.

2009 stringr: a consistent interface to R's string functions. Available from CRAN.

2009 mutatr: mutable objects with prototype based inheritance. Available from CRAN.

2009 Contributor to itertools package.

2008- tourr: an implementation of touring algorithms (grand, guided, ...) in R. With Dianne Cook and Barret Schloerke.

2008- plyr: a set of cognitive and computational tools for data analysis. Currently these tools apply to situations where you have a large problem that can be solved by breaking it into pieces, operating on each piece and then joining the results back together.

2007-8 classify: explore high-dimensional classification boundaries. Works with any classification algorithm, using a series of adaptive methods to determine class boundaries. <http://had.co.nz/classify>. Available from CRAN.

2007 clusterfly: explore clustering results in high dimensional spaces. Provides general methods for any clustering algorithm, plus tailored visualisations for hierarchical clustering, self-organising maps and model based clustering. Builds on the ideas of Andreas Buja and Dianne Cook. <http://had.co.nz/clusterfly>. Available from CRAN.

2007-9 DescribeDisplay: the R part of a system to turn exploratory GGobi graphics into publication quality R graphics. Builds on the work of Deborah F Swayne, Andreas Buja, and Dianne Cook. <http://www.ggobi.org/describe-display/>. Available from CRAN.

2007 hints: gives hints on what functions you might want to apply to an object you have created. With Sanford Weisberg. Available from CRAN.

2007-8 fda: helped James Ramsay and Bernard Silverman to turn their S code for functional data analysis into an R package. With recent contributions by Spencer Graves. <http://www.psych.mcgill.ca/misc/fda/>. Available from CRAN.

2007-9 rggobi: redevelopment of Rggobi package with Michael Lawrence to improve robustness and ease of use for casual user. Builds on the work of Duncan Temple Lang and Deborah F Swayne. <http://www.ggobi.org/rggobi>. Available from CRAN.

2006 localmds: a friendly R interface for C code written by Lisha Chen for performing local multidimensional scaling for graph layout, ordination and dimension reduction. Available from authors.

2006 lvpplot: an R package for letter value plots, an extension of the Tukey's boxplot to deal with large data. In concert with Heike Hofmann and Karen Kafadar. Available from CRAN.

2006-8 meifly: **m**odels **e**xplored **i**nteractively. Uses interactive graphics to explore ensembles of linear models. Visualises descriptive statistics on five levels: model, model-parameter, parameter, model-observation, and observation. Sole contributor, but incorporates many ideas from Andreas Buja, Duncan Temple Lang, Deborah Nolan, Philip Dixon and

Antony Unwin. Available from CRAN.

2005– reshape: flexibly reshape and aggregate data into any desired form. Built on underlying theory using two fundamental operations, melt and cast. <http://had.co.nz/reshape>. Available from CRAN.

Teaching

Short courses

2014 Jun 2014, Swedish Statistics Summer School.
Jul 2014, UseR! “An introduction to dplyr”.
Sep 2014, New York, NY. “R masterclass”.
Sep 2014, Strata. “Data manipulation with dplyr”.

2013 Jul 2013, useR 2013!. “Introduction to Rcpp”, with Romain Francois.
Jul 2013, London. “Introduction to Rcpp”, with Romain Francois.
June 2013, IARC. “Introduction to data analysis in R”.
May 2013, UBC. “Introcution to ggplot2”.
Apr 2013, UCLA. “Introduction to Rcpp”.
Mar 2013, ENAR. “Introduction to ggplot2”.
Mar 2013, Investigate Reporting and Editing Conference. “Introduction to data science”.
Feb 2013, MD Anderson. “Data analysis in R”.

2011 “R development masterclass”. Two day course. New York City, NY. December 2011.
2011 “Package development in R”. One day course. CSIRO, Canberra, Australia. October 2011.
2011 “Data visualisation in R with plyr and ggplot2”. One day course. Rice University. September 2011.
2011 “Visualising climate change”, with Di Cook and Heike Hofmann. Five day course. Reykjavik, Iceland. August 2011.
2011 “Data visualization and transformation with R packages ggplot2 and plyr”. Two day course. Ebay, Palo Alto, CA. August 2011.

2011 “R development masterclass”. Two day course. San Francisco, CA. June 2011.

2010 “Visualising data with ggplot2”. One day course. Institute of Statistical Mathematics, Tokyo, Japan. November 2010.
2010 “Data Visualisation in R: Harnessing the power of ggplot2 to produce elegant data graphics”. R Guru course sponsored by Mango Solutions, London, England. November 2010.
2010 “Creating visualisations in R”. Half-day course. VisWeek 2010, Salt Lake City, UT. October 2010.
2010 “Introduction to visualisation with R”. Two one-day courses. University of Vanderbilt, Nashville, TN. July 2010.
2010 “Visualisation and data manipulation in R”. Two day course. World Health Organisation, Geneva, Switzerland. July 2010.
2010 “Introduction to visualisation with R”. World Health Organisation, Geneva, Switzerland. 20 participants. March 2010.
2010 Data visualisation mini course at Rice University, Houston TX. 106 participants. February 2010.
2009 “Data manipulation and visualisation with R”. A two day course covering the basics of data manipulation and visualisation in R. Taught to the Biology Department of University of Missouri, St Louis, MO. 21 participants. October 2009.

- 2009 “Data visualisation with ggplot2”. A one day course covering static graphics with ggplot2. Taught at ISMI Manufacturing Week, Austin, TX. 15 participants. October 2009.
- 2009 “Looking at data”, with Dianne Cook and Heike Hofmann. A two day course covering static graphics with ggplot2 and interactive graphics with GGobi. Washington, DC. 28 participants. August 2009.
- 2009 “An introduction to plyr: understanding large data with many models”. A half-day course taught at useR! 2009. 20 participants. June 2009.
- 2007 “Statistical Graphics for High-D Data,” with Deborah F. Swayne and Dianne Cook. A half-day introduction to interactive statistical graphics for Infovis practitioners, with an emphasis on visual classification and clustering. Taught as part of Infovis 2007. Sacramento, CA. October 2007.
- 2007 “Practical data exploration with reshape, ggplot and GGobi”. A three day course taught at the Institute of Environmental Sciences, University of Zurich, Zurich, Switzerland. July 2007. <http://had.co.nz/zurich-2007>.
- 2007 “Looking at data,” with Dianne Cook. A one day introduction to static and interactive graphics using ggplot and GGobi. 10 participants. Salt Lake City, UT, July 2007. <http://lookingatdata.com/jsm-2007.html>.
- 2007 “Introduction to interactive graphics.” A half-day course taught to Novartis employees. Introduced participants to interactive graphics methodology, including linked brushing and the grand tour. Basel, Switzerland, July 2007.
- 2005–6 “R/S-plus Fundamentals and Programming Techniques” taught four two-day courses to 10–15 students on behalf of XLSolutions, a private training company. Various locations in the US. 2005–2006.

Courses

- 2012 Instructor, **Stat405**, <http://stat405.had.co.nz>. “Statistical computing and graphics”. A mix of undergraduate majors, professional masters students and PhD students (60 students). Fall 2012.
- 2012 Instructor, **Stat310**, <http://stat310.had.co.nz/>. “Mathematical statistics and probability”. A large undergraduate class (140 students). Spring 2012.
- 2011 Instructor, **Stat405**, <http://had.co.nz/stat405>. “Statistical computing and graphics”. A mix of undergraduate majors, professional masters students and PhD students (60 students). Fall 2011.
- 2011 Instructor, **Stat645**, <http://had.co.nz/stat645>. “Data visualisation”. A graduate discussion course in visualisation (10 students). Spring 2011.
- 2011 Instructor, **Stat310**, <http://had.co.nz/stat310>. “Mathematical statistics and probability”. A large undergraduate class (98 students). Spring 2011.
- 2010 Instructor, **Stat499**, <http://had.co.nz/stat499>. “Visualisation”. A discussion course in visualisation (4 students). Fall 2010.
- 2010 Instructor, **Stat405**, <http://had.co.nz/stat405>. “Statistical computing and graphics”. A mix of undergraduate majors, professional masters students and PhD students (30 students). Fall 2010.
- 2010 Instructor, **Stat491**, “Guided readings in visualisation”. A small independent study group (two students). Spring 2010.
- 2010 Instructor, **Stat310**, <http://had.co.nz/stat310>. “Mathematical statistics and probability”. A large undergraduate class (92 students). Spring 2010.
- 2009 Instructor, **Stat405**, <http://had.co.nz/stat405>. “Statistical computing and graphics”. A mix of undergraduate majors, professional masters students and PhD students (30 students). Fall 2009.
- 2009 Instructor, **Stat310**, <http://had.co.nz/stat310>. “Mathematical statistics and probability”. A large undergraduate class (60 students). Spring 2009.
- 2008 Instructor, **Stat405**, <http://had.co.nz/stat405>. “Statistical computing and graphics”. A small graduate class (8

students) teaching data analysis with R. Average student evaluation of overall course quality: 1.27 (1=Outstanding, 2=Good). Fall 2008.

2008 Instructor, **R-clinic**, <http://had.co.nz/r-clinic>. Each week, for one hour, I discussed how to “think in R”, and general strategies for solving computational problems effectively. This is not an official course, but around four faculty and ten graduate students attended each week. Spring 2008.

2007–8 Instructor, **Statistics 480**, “Applied statistical computing”, <http://had.co.nz/stat480>. Modernised curriculum to emphasise on practical data cleaning and exploration. In this 3 credit course students learned Excel, R and SAS, and had to write at least one page per week communicating their findings. Taught to a mixture of 20 undergraduate statistics majors and graduate students from other departments, Spring 2007/2008.

2007 Instructor, **Honours 322f**, “Escape from flatland”, <http://had.co.nz/hon322f>. Invited to submit proposal by Center for Excellence in Learning and Teaching. This 1 credit seminar course introduced a diverse range of students to multivariate exploratory data analysis. Each week introduced students to a new tool, and gave them a chance to experiment with it. Taught to 12 undergraduate honours students (maximum enrolment), Spring 2007.

2006 Teaching assistant, **Statistics 332**, “Visual communication of quantitative information.” Developed and taught two weeks of lectures on design and analysis of human-computer interaction experiments, to 10 advanced undergraduate and graduate students. March 2006.

2005 Relief teaching for two weeks, **Statistics 101** to 80 undergraduate students. July 2005.

2005 Relief teaching for two weeks, **Statistics 407**, “Methods of Multivariate analysis” to 15 graduate students. July 2005.

2000 Lab demonstrator, “Anatomy for Optometry Students” and “Reproduction and Development”. Faculty of Medical and Health Sciences, The University of Auckland. February–May 2000.

Workshops taught

2007 “University Teaching Seminars”, organised by the Centre for Excellence in Teaching and Learning, Iowa State University. With three other graduate students, planned and delivered a half-day series of talks for all incoming teaching assistants. Responsible for “Stand alone teaching”, with Troy Abel; a panel of international students; and “Grading: fast, fair, fun?” with Wendy Sparks. Ames, IA, August 2007.

Workshops attended

2009 “Integrating Computing into the Statistics Curricula: A Workshop to Develop Educational Materials.” Berkeley, SF. July 2009.

2009 “National Effective Teaching Institute.” Austin, TX. June 2009.

Advising

2012 Joey Huchette. “Constrained layout for visualisation”.

2012 Michael Shin. “Horizon histograms”.

2012 Kayla Schaeffer. “Quantifying gestalt principles”.

2011 James Rigby, Hyun Bin Kang and Jonathan Stewart. Supervised summer VIGRE research project extending the scatterplot to deal with large data sets.

2011 Pierre Elias. Supervised independent research project applying analytical hierarchy processing to chronic lower back pain.

2011 Henry Deng. Rice century scholar.

2010– Max Richardson. Rice ELA program.

2010–11 David Kahle. PhD committee member. Rice University, Houston, TX.

2010 Ian Fellows. Google summer of code student.

2010 Alex Gorischek. Independent study in visualisation.

2010 Gabi Quart. Independent research project.

2009 Gabi Quart, Barrett Schloerke and Dexter Gannon. Supervised summer VIGRE research project exploring the housing crisis.

2009– Garrett Grolemond. PhD advisor. Rice University, Houston, TX.

2009–11 Eric Chi. PhD committee member. Rice University, Houston, TX.

2009 Bjørn Arild Maeland. Google summer of code student.

2007–9 Barret Schloerke, Spencer Bradley and Chris Keilion. Undergraduate research assistants for GGobi grant. Iowa State University. Ames, IA.

Other

2003–4 Education technology assistant. Developed online learning and assessment resources for medical students. Faculty of Medical and Health Sciences, The University of Auckland. July 2003–July 2004.

Grants

2011–15 Investigator, “Exploring the Optimal Forecasting Frontier: How Much Room is There to Improve Subjective Forecasting Accuracy?,” IARPA, June 03 2011–June 02 2015, \$498,689, subcontract UC Berkeley.

2012 Aggregating cran statistics. Google summer of code.

2012 plyr 2.0. Google.

2010 Inference for Statistical Graphics (NSF 1007877).

2010 Google summer of code grant for developing GUI for ggplot2. Student researcher: Ian Fellows.

2010 Grant from Revolutions Computing to improve R documentation using statistics and machine learning.

2009 Grant from Becker-Dickinson improve performance of ggplot2 and plyr packages.

2009 Google summer of code grant for development of <http://crantastic.org>. Student researcher: Bjorn Arild Maelund.

2008 Travel grant for World Congress on Statistics and Probability.

2007–2010 Involved in development of NSF Cyber-Enabled Discovery and Innovation grant to fund development of biblio- and biographic data and tools to enhance discovery, data mining and distribution of statistical research. In collaboration with Jim Pitman, Michael Jordan, and Gary King. Preliminary proposal had very positive reviews, and we have been invited to submit a full proposal.

2007 Funded by the IMS to develop new navigation and classification tools for the Current Index to Statistics.

2007 Co-author of NSF grant 0706949 with D. Cook and H. Hofmann. “Statistical Graphics Research with GGobi”. Goals: develop tools and techniques for multivariate longitudinal and spatio-temporal data; advance graphics research by developing interactive faceting tools, new tour capabilities, and new methods for logical linking; build infrastructure to be used by others.

2006 Funded by NSF grant for research project with Dr David Otis, Natural Resource and Ecology. Comparing new model for mark-recapture data to existing techniques. Large scale numerical simulation, model fitting and summaries. Bootstrapping estimates of confidence intervals.

Student grants and awards

- 2012 Barret Schloerke. NSF Graduate research fellowship.
- 2011 Garrett Grolemond. NSF East Asia and Pacific summer institutes for US graduate students (EAPSI).

Internships

- 2007 Modelling and Simulation group, Novartis AG. Basel, Switzerland. Invited by Anthony Rossini. Visualisation of modelling results for high-profile study; development of, and training on, exploratory graphical methods for pharmaceutical studies.
- 2006 Department of Computer Oriented Statistics and Data Analysis, University of Augsburg. Augsburg, Germany. Invited by Antony Unwin.

Service

Statistical computing

- 2014- Member, R Foundation.
- 2013-15 Chair-elect, Chair and Past-chair, Section on Statistical Graphics, American Statistical Association.
- 2012 Program chair, Section on Statistical Graphics, American Statistical Association
- 2011 Panelist, National Science Foundation.
- 2011- Series editor, Chapman and Hall R Series.
- 2011 Publications officer, Section on Statistical Computing, American Statistical Association
- 2010 Advisory board member for StatProb, The Encyclopedia Sponsored by Statistics and Probability Societies.
- 2010- Member of young statisticians advisory board to the ISI.
- 2010 “Visualising data”. Coffee roundtable, Section on Quality and Productivity, JSM 2010.
- 2010 Organised invited session for JSM 2010 on “The quantified self: personal data collection, analysis and exploration.”
- 2010-11 Member of program committee for useR! 2011.
- 2010 Book reviewer for Springer.
- 2009 “Patterns of Data Manipulation”. Coffee roundtable, Section for Statistical Programmers and Analysts, JSM 2009.
- 2009 Organised ASA Statistical Computing and Graphics Data Expo 2009, <http://stat-computing.org/dataexpo/2009/>. Session chair for associated poster session at JSM 2009, with 14 posters submitted.
- 2009-12 Judge for John M. Chambers Statistical Software Award.
- 2008- Development of crantastic.org, a community site for finding, rating and reviewing R packages.
- 2008-14 Management committee, Current Index to Statistics (IMS representative).
- 2007 Founding Secretary, GGobi Foundation.
- 2007 Member of program committee, web site developer, and session chair for useR! conference. <http://user2007.org>.
- 2006- Member of IMS Electronic Issues committee.
- 2006-7 Web developer for the Department of Statistics at Iowa State University.
- 2005-7 Member of computation committee. Department of Statistics, Iowa State University.

- 2005–7 Web master for Staters statistics student organisation.
- 2006– Web master, ASA Sections on Statistical Computing and Graphics.

Journals

- 2015–17 Member, JCGS management committee.
- 2014– Associate Editor, The American Statistician.
- 2013 Editor-in-chief, The R Journal.
- 2011– Associate Editor, Journal of Computational and Graphical Statistics.
- 2011–12 Associate Editor, The R Journal.
- 2007– Associate Editor, Journal of Statistical Software.

Conferences

- 2015 Program committee, UseR.
- 2012 Program chair, Interface.
- 2011–13 Program committee, Infovis.

University

- 2011–12 Chair, Undergraduate Committee. Department of Statistics.
- 2010 Member of Senate Committee on Teaching.
- 2009–12 Member of Rice Center for Engineering Leadership Internal Advisory Committee.
- 2009–10 Faculty sponsor for Rice Women's soccer team.
- 2009–12 Divisional advisor for McMurry college.
- 2008–12 Major advisor for Statistics.
- 2008–12 Member of Oshman Engineering Design Kitchen Faculty Committee.

Local

- 2011 Chair of Young Researchers session at the Lehmann Symposium, Rice University, 2011.
- 2008–10 President-elect, President and Past-president of Houston Area Chapter of the American Statistical Association.

Other

- 2005 Organised VIGRE seminar series on statistical graphics. This involved arranging a weekly speaker or discussion topic for two semesters. Department of Statistics, Iowa State University.
- 2004–6 Statistical consultant. Provided consulting to over 30 PhD students and staff from over ten departments. Used both SAS and R extensively. Department of Statistics, Iowa State University.

Advisory roles

- 2012– Advisor, Metamarkets.
- 2011–13 Advisor, Luckysort.
- 2011– Advisor, rOpenSci.