NICK STRAYER

I have made visualizations viewed by hundreds of thousands of people[1], sped up query times for 25 terabytes of data by an average of 4,800 times[2], and built packages for R[3] that let you do magic[4].
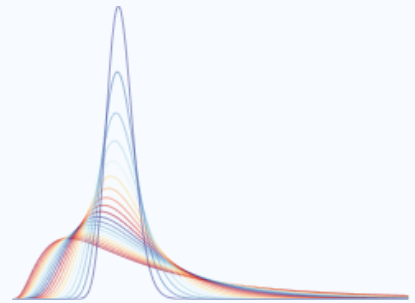
Currently searching for a data science position that allows me to build tools using visualization and machine learning to help people explore and understand their data.

🎓 EDUCATION

PhD. Candidate, Biostatistics

Vanderbilt University   📍 Nashville, TN

2020 | 2015

- Working on Bayesian network models & interactive visualization platforms
- University Graduate Fellow

B.S., Mathematics, Statistics (minor C.S.)

University of Vermont   📍 Burlington, VT

2015 | 2011

- Thesis: An agent based model of Diel Vertical Migration patterns of Mysis diluviana

💻 RESEARCH EXPERIENCE

Graduate Research Assistant

TBILab (Yaomin Xu's Lab)   📍 Vanderbilt University

2020 | 2015

- Primarily working with large EHR and Biobank datasets.
- Developing network-based methods to investigate and visualize clinically relevant patterns in data.

Data Science Researcher

Data Science Lab   📍 Johns Hopkins University

2018 | 2017

- Building R Shiny applications in the contexts of wearables and statistics education.
- Work primarily done in R Shiny and Javascript (node and d3js).

Undergraduate Researcher

Rubenstein Ecosystems Science Laboratory   📍 University of Vermont

2015 | 2013

- Analyzed and visualized data for CATOS fish tracking project.
- Head of data mining project to establish temporal trends in population densities of Mysis diluviana (Mysis).
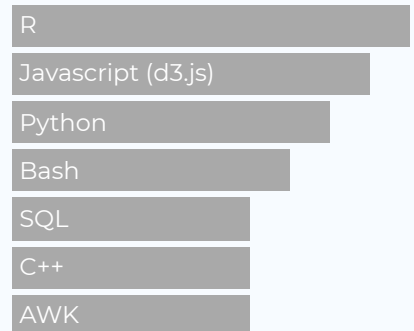- Ran project to mathematically model the migration patterns of Mysis (honors thesis project.)

View this CV online with links at *nickstrayer.me/cv*

CONTACT

- ✉ nick.strayer@gmail.com
- 🐦 NicholasStrayer
- ⌂ github.com/nstrayer
- 🔗 nickstrayer.me
- 📞 (734) 645-0110

LANGUAGE SKILLS

R

Javascript (d3.js)

Python

Bash

SQL

C++

AWK

Made with the R package page-down.

The source code is available at github.com/nstrayer/cv.

Last updated on 2019-10-11.

2015 ● **Human Computer Interaction Researcher**
LabInTheWild (Reineke Lab)   📍 University of Michigan

· Led development and implementation of interactive data visualizations
  to help users compare themselves to other demographics.

2014
 | ● **Undergraduate Researcher**
2013 Bentil Laboratory   📍 University of Vermont

· Developed mathematical model to predict the transport of sulfur
  through the environment with applications in waste cleanup.

2013
 | ● **Research Assistant**
2012 Adair Laboratory   📍 University of Vermont

· Independently analyzed and constructed statistical models for large
  data sets pertaining to carbon decomposition rates.

🧳 INDUSTRY EXPERIENCE

2016 ● **Data Journalist – Graphics Department**
New York Times   📍 New York, New York

· Reporter with the graphics desk covering topics in science, politics, and
  sport.
· Work primarily done in R, Javascript, and Adobe Illustrator.

2015 ● **Engineering Intern – User Experience**
Dealer.com   📍 Burlington, VT

· Built internal tool to help analyze and visualize user interaction with
  back-end products.

2015 ● **Data Science Intern**
Dealer.com   📍 Burlington, VT

· Worked with the product analytics team to help parse and visualize
  large stores of data to drive business decisions.

2015
 | ● **Data Artist In Residence**
2014 Conduce   📍 Carpinteria, CA

· Envisioned, prototyped and implemented visualization framework in the
  course of one month.
· Constructed training protocol for bringing third parties up to speed with
  new protocol.

2014 ● **Software Engineering Intern**
Conduce   📍 Carpinteria, CA

· Incorporated d3.js to the company's main software platform.

I have worked in a variety of roles ranging from journalist to software engineer to data scientist. I like collaborative environments where I can learn from my peers.

⊡ TEACHING EXPERIENCE

2019

Data Visualization Best Practices
DataCamp

- Designed from bottom up course to teach best practices for scientific visualizations.
- Uses R and ggplot2.
- In top 10% on platform by popularity.

2019

Improving your visualization in Python
DataCamp

- Designed from bottom up course to teach advanced methods for enhancing visualization.
- Uses python, matplotlib, and seaborn.

2018 | 2017

Advanced Statistical Learning and Inference
Vanderbilt Biostatistics Department   📍 Nashville, TN

- TA and lectured
- Topics covered from penalized regression to boosted trees and neural networks
- Highest level course offered in department

2018

Advanced Statistical Computing
Vanderbilt Biostatistics Department   📍 Nashville, TN

- TA and lectured
- Covered modern statistical computing algorithms
- 4th year PhD level class

2017

Statistical Computing in R
Vanderbilt Biostatistics Department   📍 Nashville, TN

- TA and lectured
- Covered introduction to R language for statistics applications
- Graduate level class

I am passionate about education. I believe that no topic is too complex if the teacher is empathetic and willing to think about new methods of approaching task.

📈 SELECTED DATA SCIENCE WRITING

2019

Using AWK and R to Parse 25tb[5]
LiveFreeOrDichotomize.com

- Story of parsing large amounts of genomics data.
- Provided advice for dealing with data much larger than disk.
- Reached top of HackerNews.

2018

Classifying physical activity from smartphone data[6]
RStudio Tensorflow Blog

- Walk through of training a convolutional neural network to achieve state of the art recognition of activities from accelerometer data.
- Contracted article.

I regularly blog about data science and visualization on my blog LiveFreeOrDichotomize.

2018 ● **The United States of Seasons**[7]

LiveFreeOrDichotomize.com

· GIS analysis of weather data to find the most 'seasonal' locations in United States
· Used Bayesian regression methods for smoothing sparse geospatial data.

2017 ● **A year as told by fitbit**[8]

LiveFreeOrDichotomize.com

· Analyzing a full years worth of second-level heart rate data from wearable device.
· Demonstrated visualization-based inference for large data.

2017 ● **MCMC and the case of the spilled seeds**[9]

LiveFreeOrDichotomize.com

· Full Bayesian MCMC sampler running in your browser.
· Coded from scratch in vanilla Javascript.

2017 ● **The Traveling Metallurgist**[10]

LiveFreeOrDichotomize.com

· Pure javascript implementation of traveling salesman solution using simulated annealing.
· Allows reader to customize the number and location of cities to attempt to trick the algorithm.

SELECTED PRESS (ABOUT)

2017 ● **Great paper? Swipe right on the new 'Tinder for preprints' app**[11]

Science

· Story of the app Papr[12] made with Jeff Leek and Lucy D'Agostino McGowan.

2017 ● **Swipe right for science: Papr app is 'Tinder for preprints'**[13]

Nature News

· Second press article for app Papr.

2016 ● **The Deeper Story in the Data**[14]

University of Vermont Quarterly

· Story on my path post graduation and the power of narrative.

SELECTED PRESS (BY)

2016 ● **The Great Student Migration**[15]

The New York Times

· Most shared and discussed article from the New York Times for August 2016.

2016 ● **Wildfires are Getting Worse, The New York Times**[16]

The New York Times

· GIS analysis and modeling of fire patterns and trends
· Data in collaboration with NASA and USGS

2016 ● **Who's Speaking at the Democratic National Convention?**[17]

The New York Times

· Data scraped from CSPAN records to figure out who talked and past conventions.

2016 ● **Who's Speaking at the Republican National Convention?**[18]

The New York Times

· Used same data scraping techniques as Who's Speaking at the Democratic National Convention?

2016 ● **A Trail of Terror in Nice, Block by Block**[19]

The New York Times

· Led research effort to put together story of 2016 terrorist attack in Nice, France in less than 12 hours.
· Work won Silver medal at Malofiej 2017, and gold at Society of News and Design.

▤ SELECTED PUBLICATIONS, POSTERS, AND TALKS

2019 ● **Charge Reductions Associated with Shortening Time to Recovery in Septic Shock**[20]

Chest

· Authored with Wesley H. Self, MD MPH; Dandan Liu, PhD; Stephan Russ, MD, MPH; Michael J. Ward, MD, PhD, MBA; Nathan I. Shapiro, MD, MPH; Todd W. Rice, MD, MSc; Matthew W. Semler, MD, MSc.

2019 ● **Multimorbidity Explorer | A shiny app for exploring EHR and biobank data**[21]

RStudio::conf 2019

· Contributed Poster. Authored with Yaomin Xu.

2019 ● **Taking a network view of EHR and Biobank data to find explainable multivariate patterns**[22]

Vanderbilt Biostatistics Seminar Series

· University wide seminar series.

2019 ● **Patient-specific risk factors independently influence survival in Myelodysplastic Syndromes in an unbiased review of EHR records**

Under-Review (copy available upon request.)

· Bayesian network analysis used to find novel subgroups of patients with Myelodysplastic Syndromes (MDS).
· Analysis done using method built for my dissertation.

2019 ● **Patient specific comorbidities impact overall survival in myelofibrosis**

Under-Review (copy available upon request.)

· Bayesian network analysis used to find robust novel subgroups of patients with given genetic mutations.
· Analysis done using method built for my dissertation.

2018 ● **R timelineViz: Visualizing the distribution of study events in longitudinal studies**

Under-Review (copy available upon request.)

· Authored with Alex Sunderman of the Vanderbilt Department of Epidemiology.

2017 ● **Continuous Classification using Deep Neural Networks**[23]

Vanderbilt Biostatistics Qualification Exam

· Review of methods for classifying continuous data streams using neural networks
· Successfully met qualifying examination standards

2015 ● **Asymmetric Linkage Disequilibrium: Tools for Dissecting Multiallelic LD**

Journal of Human Immunology

· Authored with Richard Single, Vanja Paunic, Mark Albrecht, and Martin Maiers.

2015 ● **An Agent Based Model of Mysis Migration**[24]

International Association of Great Lakes Research Conference

· Authored with Brian O'Malley, Sture Hansson, and Jason Stockwell.

2015 ● **Declines of Mysis diluviana in the Great Lakes**

Journal of Great Lakes Research

· Authored with Peter Euclide and Jason Stockwell.

🔗 LINKS

1: https://www.nytimes.com/interactive/2016/08/26/us/college-student-migration.html
2: https://livefreeordichotomize.com/2019/06/04/using_awk_and_r_to_parse_25tb/
3: https://github.com/nstrayer/shinysense
4: http://nickstrayer.me/dataDayTexas/
5: https://livefreeordichotomize.com/2019/06/04/using_awk_and_r_to_parse_25tb/
6: https://blogs.rstudio.com/tensorflow/posts/2018-07-17-activity-detection/
7: https://livefreeordichotomize.com/2018/02/12/the-united-states-of-seasons/
8: https://livefreeordichotomize.com/2017/12/27/a-year-as-told-by-fitbit/
9: https://livefreeordichotomize.com/2017/10/14/mcmc-and-the-case-of-the-spilled-seeds/
10: https://livefreeordichotomize.com/2017/09/25/the-traveling-metallurgist/
11: https://www.sciencemag.org/news/2017/06/great-paper-swipe-right-new-tinder-preprints-app

12: https://jhubiostatistics.shinyapps.io/papr/

13: https://www.nature.com/news/swipe-right-for-science-papr-app-is-tinder-for-pre-prints-1.22163

14: https://www.uvm.edu/uvmnews/news/deeper-story-data

15: https://www.nytimes.com/interactive/2016/08/26/us/college-student-migration.html?smid=pl-share

16: https://www.nytimes.com/interactive/2016/07/25/us/wildfire-seasons-los-angeles.html

17: https://www.nytimes.com/2016/07/26/upshot/democrats-may-not-be-unified-but-their-convention-speakers-are.html

18: https://www.nytimes.com/2016/07/19/upshot/whos-not-speaking-how-this-republican-convention-differs.html?smid=pl-share

19: https://www.nytimes.com/interactive/2016/07/14/world/europe/trail-of-terror-france.html

20: https://www.ncbi.nlm.nih.gov/pubmed/30419234

21: http://nickstrayer.me/rstudioconf19_me-poster/

22: http://nickstrayer.me/biostat_seminar/

23: http://nickstrayer.me/qualifying_exam/

24: https://www.semanticscholar.org/paper/An-Agent-Based-Model-of-the-Diel-Vertical-Migration-Strayer-Stockwell/40493c78e8ecf22bd882d17ec99fd913ec4b9820