Jasmine Dumas

Data Scientist

jasmine.dumas@gmail.com | jasdumas.github.io | Connecticut, USA

About

I'm seeking to utilize my computational programming & data analysis skills in a position that focuses in developing innovative open source software tools. I value hard work, creativity and openness. I'm an active member of the R Programming community and have contributed by formerly co-hosting a podcast on the language, RTalk and as a organization member & curator of the online newsletter, RWeekly.

Featured Projects:

- Web scraping for craft brewery ratings from Beer Advocate with R and import.io
- Shiny app for gene expression analysis for bioinformaticians
- Dashboard app of Reported Crimes in Hartford Connecticut

R Packages

- `ttbbeer`: R data package of beer statistics from U.S. Department of the Treasury, Alcohol and Tobacco Tax and Trade Bureau (TTB) | CRAN Link | GitHub Link
- `dumas`: My personal R package filled with useful functions for data analysis | GitHub Link
- `shinyGEO`: Gene Expression Omnibus Analysis with Shiny | GitHub Link | Web App

Skills

- **R Programming**: ggplot2, Shiny, dplyr, reshape2, plotly, Bioconductor, knitr, RMarkdown, R package development
- **Statistical Software Packages**: Minitab, SPSS, SAS
- **Machine Learning**: Regression Analysis, GLM, GBM, xgboost, Clustering, Random Forest, Decision Trees, Ensemble Modeling
- **Python**: SQLite3, NumPy, pandas, scikit-learn
- **Database Processing**: basic SQL queries, Database Management (SQLDeveloper, Oracle)
- **Web Development**: HTML, Bootstrap Web Framework
- **General Computing**: GitHub, Git, version control, Markdown, Unix/Command line scripting

Education

- **DePaul University**, (2015 - Expected Graduation 2018)
  - **Master of Science Candidate in Predictive Analytics**
  - Coursework: Intro to Programming with Python, Data Analytics & Regression, Database Processing for Large-scale Analytics, Advanced Data Analysis, Knowledge Discovery Technologies
- Additional Graduate Coursework: Johns Hopkins University, Engineering For Professionals: Data Structures, Computer Organization (Summer 2016)

- **University of Hartford**, (2008 - 2012)
  - **Bachelor of Science in Biomedical Engineering**
  - Coursework: Engineering Computer Applications, Calculus (1-2, Multivariable), Differential Equations, Independent Research, Engineering Design, Statics, Dynamics, Mechanics of Materials, Bio-materials

Data Science Experience

- **The Hartford Financial Services (NYSE: HIG)**
  - **Associate Data Scientist**, Predictive Analytics & Research - Auto Team, (April 2016 - Present)
  - Research & Implementation of **machine learning techniques** in **variable reduction and selection** to create predictive models for auto insurance class plans that improve loss ratio estimates and drive strategic pricing changes
  - **Competitive analysis** modeling to understand features that contribute to premium differences
  - "Unconstrained" Loss modeling for +700 variable selections
  - Developed an **internal R package** for the Personal Lines department
  - R/Shiny application development for model monitoring and diagnostic business intelligence tools
  - **Data Science Intern**, (Nov. 2015 - March 2016)
  - Developing a model monitoring web application with **R/Shiny** for auto insurance predictive model monitoring
  - Diagnostic residual analysis modeling to evaluate loss ratio in adverse risk segments for class plan changes

- **University of Connecticut Institute for Systems Genomics (UCONN)**, (Sept. 2015 - January 2016)
  - **Bioinformatics Intern**
  - Computational Programming and Unix scripting to contribute to the annotation of the Douglas-fir & Walnut genome

- **Google Summer of Code**, (May 2015 - Aug. 2015)
  - **Student Developer** for the R Project for Statistical Computing Organization
  - Developed an web application with R & Shiny to automate differential expression and survival analysis of microarray gene expression datasets from the NIH Gene Expression Omnibus github.com/jasdumas/shinyGEO

Peer-Reviewed Research Papers

- **Dumas J**, Gargano MA, Dancik GM. shinyGEO: a web-based application for analyzing Gene Expression Omnibus datasets. Bioinformatics. 2016 Aug 8. doi:10.1093/bioinformatics/btw519 | Paper link

- **Dumas J**, et.al., "Feasibility of an electronic stethoscope system for monitoring neonatal bowel sounds." Connecticut Medicine, Volume 77, Number 8, pp. 467-471, September 2013. | bit.ly/JMD-connmed

Posters & Presentations

- IPA & data frames: Open Government Data and Beer Analytics | Presented Virtually to Cardiff, UK R User group, August 2016.
- useR! 2016 Conference International Annual Meeting (Stanford, CA): shinyGEO: An online tool for biomarker analysis in Gene Expression Omnibus (GEO) datasets
- American Association for Cancer Research Annual Meeting (New Orleans, LA) 2016: *An online tool for biomarker analysis in Gene Expression Omnibus (GEO) datasets*
- Northeast American Society of Engineering Education Conference (Lowell, MA) 2012: Dual Presenter - bit.ly/JMD-asee1, bit.ly/JMD-asee2
- Neonatal/Perinatal Research Symposium at Connecticut Children's Medical Center, (Hartford, CT) 2012

Honors & Awards

- useR! 2016 Conference Diversity Scholarship
- 2015 Google Summer of Code Program Stipend for the R Project for Statistical Computing
- 2012 NASA - Connecticut Space Grant Award: Senior Design Project Funding
- University of Hartford: Academic President's List, Dean's List