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.

Featured Projects:
I have experience using my data science toolkit in several projects which include:

- web scraping for craft brewery ratings from Beer Advocate
- developing a gene expression analysis web application for bioinformaticians
- exploratory data analysis on insurance company complaints from the CT open data portal

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, GitHub
- **dumas**: My personal R package filled with useful functions for data analysis. GitHub
- **shinyGEO**: Gene Expression Omnibus Analysis with Shiny. GitHub, web application

Skills

- R Programming (ggplot2, Shiny, dplyr, reshape2, plotly, Bioconductor, knitr, Rmarkdown)
- Statistical Software Packages: Minitab, SPSS, Towers Watson - Emblem, SAS
- Python (SQLite3, NumPy, pandas, scikit-learn)
- Database Processing: SQL (SQLDeveloper, Oracle)
- Web Development: HTML, Bootstrap
- General Computing: GitHub, Git, version control, Markdown, Mac, Windows, Unix/Command line scripting

Education

- Graduate Coursework: Johns Hopkins University, Engineering For Professionals: Data Structures, Computer Organization (Summer 2016)
- Graduate Coursework at DePaul University as an M.S. Candidate: Intro to Programming, Data Analytics & Regression, Database Processing for Large-scale Analytics, Advanced Data Analysis, Knowledge Discovery Technologies (2014 - Present)

- University of California, Irvine Extension, (2013 - 2013)
  - **Professional Certificate in Medical Product Development**

  - Coursework: Medical Product Life-Cycle Management, Regulatory Requirements for Medical Devices, Medical Product Quality Systems, Medical Device Design and Evaluation, Fundamentals of Clinical Trials

---

- University of Hartford, (2008 - 2012)
  - **Bachelor of Science in Biomedical Engineering**

  - Coursework: Engineering Computer Applications, Calculus (1-2, Multivariable), Differential Equations, Research, Engineering Design, Statics, Dynamics, Mechacnics of Materials, Biomaterials

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.
- 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.

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
- 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

Engineering Experience

Medtronic (NYSE: MDT), (May 2014 - Sept. 2015)
R&D Engineering Technician III
- Statistical analysis & visualization of mechanical product testing data of surgical instruments to provide insight into product performance and evaluation for FDA submissions
- **Major Accomplishment**: Technical Forum Analytics member uncovering insights in meeting attendance with R scripts

---

QDx, Inc. / Abbott Point-of-Care (NYSE: ABT), (June 2013 - Jan. 2014)
R&D Engineering Intern
- Data Collection, Statistical analysis, and Data Mining of animal & human specimens CBC (complete blood count) results of an image based, reagent-free hematology analyzer in preparation for clinical trials at a start-up
- **Major Accomplishment**: Designed an experiment leading to a novel feature enhancement for anticoagulants

AdChem Manufacturing Technologies, Inc., (May 2012 - May 2013)
Manufacturing Engineer
- Developed over 250 manufacturing process methods for complex commercial and military jet engine sheet metal brackets for Pratt & Whitney, using CNC Metrology Software, 3D printing & Solidworks CAD - **Major Accomplishment**: Developed a new manufacturing method & CAD Model for part identification

Research

Research Interests: Bioinformatics software, Reproducible research and analysis, Open Government Data, Computational programming, Beer analytics, Webscraping

Peer-Reviewed Research Papers

- **Dumas, J.**, Dancik, GM., "shinyGEO: a web-based tool for downloading, cleaning, & modeling of microarray gene expression datasets" [Pending Publication in the Bioinformatics Journal]
- **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

- 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

Community Involvement

- Co-host of **R Talk**: A podcast about the R Programming language rtalk.org
- NumFOCUS, Community Member
- Bioinformatics Organization, Associate Member
- Python Software Foundation, Basic Member
- Connecticut Predictive Analytics, Meetup.com Group Member
- Epsilon Pi Tau, Honor Society, General Member Inducted in 2012