Patrick McCarthy

http://asteriske.github.io
patrick.jmc@gmail.com
(347) 329-3638

SKILLS

- R, Python Scikit-Learn, Pandas, SAS
- AWS, Spark, git, SVN, Netezza
- Machine learning, Survey Sampling

- Java, C++, bash (sed, awk, grep)
- Linux, OS X, Windows
- Data visualization in D3.js, Django

EXPERIENCE

Data Scientist, eXelate, Inc. (2013-Present, acquired by The Nielsen Company 2015)
- Designed and built end-to-end modeling pipelines against 1bn+ records for use in secure remote environments
- Built several hundred real-time scoring models from multi-gigabyte sparse-feature datasets, including many where $p >> n$
- Implemented high-performance concurrent ETL system Python, EC2 and S3
- Researched novel applications of statistical practice to business problems, including latent class analysis, Bayesian models
- Developed process automation and dashboard applications in Python, R, d3.js

Campaign Analysis, Eric Garcetti for Mayor of Los Angeles (May 2013)
Consultant. Mined donor database to uncover trends, discovered potential repeat donors. Communicated results, recommendations to non-quantitative staff. http://goo.gl/SDERNa

LA Sheriff's Office Corruption Investigation, Witness LA (September 2012)
Consultant. Support for investigative print journalism. Generated usable dataset for analysis from personnel and financial records, discovered high likelihood of "pay for play" promotion practices. *Best Database-Driven Journalism award, LA Press Club*. http://goo.gl/GYvI7R

Quantitative Analyst Intern (June 2012 – September 2012)
Acumen LLC, Burlingame CA
- Enhanced client relationships by developing a general use JavaScript visualization tool to intuitively display hierarchical datasets http://goo.gl/iO6BcY
- Primary research for Medicare Administration /Affordable Care Act contract
- Developed new healthcare policy insights through exploration and analysis of a multi-gigabyte dataset, including extensive use of SAS macros and ODS

Technical Specialist (September 2008 – September 2011)
Information Technology, Boston University Medical Campus, Boston MA
- Managed teams through multiple phases of enterprise infrastructure deployment including research, initial implementation and ongoing maintenance
- Took initiative in streamlining department operations, both in procedure and technology

EDUCATION

M.S. Statistics, June 2013
University of California Los Angeles, Los Angeles CA
Thesis: *A Study on Conditional Likelihood Estimation for Survey Sampling*

B.A. Economics and Mathematics, May 2008
Boston University, Boston MA