

Erin Shellman

937-321-1129
erin.shellman@gmail.com
www.erinshellman.com

I'm a data scientist with experience in a broad range of industries including retail, cloud computing, and biotechnology. Along the way, I've built recommendation engines, web scrapers, interactive visualizations, and analyzed terabytes of data. I love to teach and speak and do so often through talks, as co-organizer of PyLadies-Seattle, and as an instructor at the University of Washington's Professional and Continuing Education program.

Skill Summary: Python, R, Spark, SQL/Hive, statistical analysis, machine learning, teaching, technical writing

edu

University of Michigan
Ann Arbor, MI

PhD - Bioinformatics *class of 2012*
MS - Biostatistics *class of 2008*

Case Western Reserve University
Cleveland, OH

BA - Economics, Evolutionary Biology
class of 2006

work

Senior Data Scientist, Zymergen
Seattle, WA – November 2015 - Present

Zymergen uses genomic technology and engineering principles to do end-to-end microbial strain improvement. At Zymergen, I

- support our high-throughput screening platform with analytical automation
- use machine learning techniques to make predictive strain designs
- developed communication and leadership skills through career coaching

Instructor - University of Washington PCE
Seattle, WA – March 2015 - Present

I teach a 10-week data mining course as part of the business intelligence certification program. This is a project-based course focusing on prediction of numerical outcomes, classification, and unsupervised learning. I've open-sourced my course materials which can be found here: <http://bit.ly/1P2mbxS>

Research Scientist II, Amazon Web Services - S3
Seattle, WA – March 2015 - October 2015

S3 is a cloud-based storage service. During my tenure, it served 1.5 million requests per second and generated unprecedented volumes of data. At S3, I

- processed terabytes of web server log data with Scala + Spark using the AWS ecosystem including EMR, S3, EC2, and RedShift
- researched customer usage patterns to inform business and engineering decisions

Data Scientist II, Nordstrom Data Lab
Seattle, WA – March 2013 - March 2015

The Data Lab is a small, collaborative team that builds data-driven services. As part of the lab, I

- wrote recommendation algorithms in Python, including recommendations based on real-time customer behavior
- built a personalized email campaign to remind customers to replenish beauty products based on individual consumption patterns
- mentored an intern

community

Promoting women and girls in technology is my passion. PyLadies is an international mentorship group for women who program in Python and I've been an organizer for the Seattle chapter for nearly two years. Our chapter is very active and hosts weekly hack-nights and monthly talk nights in addition to attending conferences as a community sponsor.



Additionally, I taught an after school coding club for girls at Ingraham High School through the Girls Who Code national organization from 2014 - 2016. I taught computer science fundamentals and coding in Scratch and Python.

publications

Erin R. Shellman, Yu Chen, Xiaoxia Lin, Charles F. Burant, Santiago Schnell. *Metabolic network motifs can provide novel insights into evolution: The evolutionary origin of Eukaryotic organelles as a case study*, Computational Biology and Chemistry, 53PB, 242–250. December 2014. <http://bit.ly/1MNFYlQ>

Erin R. Shellman, Charles F. Burant and Santiago Schnell. *Network motifs provide signatures that characterize metabolism*, Molecular BioSystems, 9(3), 352–360. December 2012. <http://rsc.li/1Pu4BSW>

Ali Izadpanah, R. James Barnard, Anna Jane E. Almeda, Gayle C. Baldwin, Sydney A. Bridges, Erin R. Shellman, Charles F. Burant, Christian K. Roberts. *A short-term diet and exercise intervention ameliorates inflammation and markers of metabolic health in overweight/obese children*. American Journal of Physiology - Endocrinology and Metabolism, 303(4), E542–50. August 2012. <http://bit.ly/1Ed07i0>

talks

Catching the most with high-throughput screening.
Seattle Data Analytics and Machine Learning. Seattle, WA. July 2016. <http://bit.ly/2bhNUSg>

Developing Effective Data Scientists.
Invited talk at the Michigan Institute for Data Science Kick-off event. Ann Arbor, MI. Oct 2015. <http://bit.ly/1MWh2Hg>

Bot or Not?: an end-to-end data analysis in Python.
PyData. Seattle, WA. Aug 2015. <http://bit.ly/1MNH7zc>

How Nordstrom Utilizes Human Intelligence to Blend Brick-and-Mortar with Online Commerce.
Strata + Hadoop World Conference. New York, NY. Oct. 2013 & Data Science Innovation Summit. San Diego, CA. Feb 2014. <http://oreil.ly/2aS7v7x>

media

Becoming a data scientist podcast. March 2016. <http://bit.ly/2b5fEor>

Data Scientists at Work. Interview in chapter Dec 2014. <http://amzn.to/1Ed2hPy>

Interview in Dashing D3. 2013. <http://bit.ly/1KeVwha>

for more detail see: <https://www.linkedin.com/in/erinshellman>