Samantha D. Sifleet

| sdsifleet@gmail.com | Seattle WA 98117 |

Research Data Scientist with over a decade of modeling experience: 9 years in R and 7 in Python.

PROFESSIONAL EXPERIENCE:

Research Data Scientist, Alexa Skills @ Amazon (*September 2017 – present*)

- Designed, built, and implemented an automated A/B testing pipeline for social marketing that cut manual curation efforts in half and drove a 50% increase in customer engagement. Automated causal lift analysis at the campaign rather than lifecycle level.
- Built a predictive model of skill marketability at launch – prior to customer engagement – using logistic regression and Boruta automated feature selection. Created a pipeline of new content for all consumer marketing channels minimizing manual discovery effort: delivered a 50% increase from the previous catalog of marketable skills. Distilled model features into insights for developer education and marketing.
- Built multiple models measuring the causal impact of programs using the synthetic control method with elastic net feature selection and regularization. The outputs of these models continue to drive performance-based investment strategies including $8MM investment in developer incentives programs and hackathons and $250K in TV advertisements.
- Developed and taught new curriculum content for Amazon Machine Learning University focused on experimental design, power analysis, and regression. Mentored Business Analysts, Business Intelligence Engineers, and Software Development engineers with interest in the Data Science field.
- *Python (numpy, scipy, scikit-learn, pandas), R (Shiny, ggplot2, data.table etc…), Tableau, Linux and Windows OS, AWS (Redshift, S3, EC2, EMR, Athena, QuickSight, Glue), Bash*

Environmental Data Scientist, Amazon (*December 2016 – September 2017*)

- Built a predictive model to differentiate growth in spending amongst new vs. existing suppliers for $1 billion cost stream. Enabling the capture of $400 million in annual operational inefficiency.
- Designed, provisioned, and administered the data pipelines and architecture supporting the cloud computing needs of a Research Science Team within the Worldwide Sustainability Organization.
- Enabled global Economic Input-Output Life Cycle Assessment (EIO-LCA) of carbon within the Amazon supply chain and customer behavior experimentation.
- *Python (numpy, scipy, pandas), R (Shiny, ggplot2, data.table, etc…), Tableau, Linux and Windows OS, AWS (Redshift, S3, EC2, VPC/networking, IAM, KMS, Athena, DMS, QuickSight)*

Environmental Data Scientist, RTI International *(October 2014 – Dec 2016)*

- Designed, built, and deployed the Water Quality Portal (WQP) Data Discovery Tool for the US EPA. This desktop application allows users to query, clean, and visualize data from the WQP using web services. The tool enables US States and Tribes to meet reporting requirements for federal funding. Code available on github.
- Designed, built, and deployed the Toxics and Climate Change web-based application. This application enables users to visually explore and map toxic chemical releases with the changing climate variables such as sea level rise (NOAA REST webservices) and extreme heat (CDC webservices). Leader of this marketing effort to showcase data visualization, tool development, and deployment capabilities.
- Technical Lead for Data Management and Programming Support tasked with data management, data cleaning, and model implementation for the Food and Drug Administration's (FDA) Risk Ranking Model for Product Tracing (RRM-PT). Leveraged complex weighted data from the National Health and Nutrition Examination Survey (NHANES) in R. Automated a legacy version of the RRM-PT model using SAS. Developed, maintained and updated a clean database underlying the RRM-PT model. Updated the automated RRM-PT model in SAS based on input from Food Science experts. Designed and led the implementation of a SQL Server data management system.
- Designed and built a flexible multi-criteria decision support tool enabling quantitative food safety risk assessment support by the FDA. This tool includes a module to run Monte Carlo simulation for uncertainty analyses and incremental sensitivity analyses.
- Lead Developer for the EPA Renewable Fuel Standard website in the Drupal environment using interactive data visualization in multiple java script libraries including DataTables plugin for jQuery and Highcharts.

- Designed and built a desktop application that allows users to explore data from the Food and Drug Administration's (FDA) Total Diet Study (TDS). The TDS is an ongoing effort to assess the levels of chemical contaminants in American foods. This tool allows the user to view interactive graphs of contaminant levels in individual foods over time. Users may filter data by contaminant, food, and year.
- Mentored 2 Data Science interns and a full time Environmental Engineer interested in moving into a Data Science role.
- *Python (numpy, scipy, arcpy, pandas), R (Shiny, ggplot2, data.table, etc…), Javascript (D3.js, Highcharts.js, DataTables, Leaflet), HTML, CSS, MATLAB, SQL, SAS, STATA, ArcGIS, Linux and Windows OS, Digital Ocean and AWS (Redshift, S3, EC2)*

GIS and Geospatial Modeling Fellow to the EPA EnviroAtlas Project *(May 2011 – Sept 2014)*

- Designed, developed, and automated cloud computing solutions for computationally intense and spatially explicit ecosystem service models of supply and demand.
- Automated the creation of alternate future landscape data sets under various climate, population growth, and land use scenarios.
- Developed an ecosystem rarity metric and downloadable Geographic Information Systems (GIS) toolbox combining disparate data sets. The metric is designed for targeting conservation efforts and in guiding land management and development decisions.
- Served as research liaison between the EnviroAtlas and regional, federal, and international organizations including the Commission for Environmental Cooperation (CEC), the National Ecosystem Partnership (NESP), and Landscape Conservation Cooperatives (LCC).
- Recognized expert in the field of blue carbon (i.e., carbon stored in coastal wetlands).
- *Python (numpy, scipy, arcpy, pandas), R (ggplot2, data.table), MATLAB, SAS, ArcGIS, Linux and Windows OS, AWS (EC2)*

Consultant to the Commission for Environmental Cooperation *(Nov. 2012 to May 2013)*

- Contracted as an expert in Blue Carbon to conduct a Scoping Study.

Policy Research Associate at the Nicholas Institute for Environmental Policy Solutions at Duke University *(May 2010 to May 2011)*

- Conducted research on carbon sequestration in coastal ecosystems. Aggregated data and built economic models examining potential off-set payments as a means for coastal habitat protection. Delivered recommendations to the United Nations Framework Convention on Climate Change Conference of the Parties 16 (COP 16), National Ocean and Atmosphere Association (NOAA), and the World Bank.
- Managed and analyzed output from the Forest and Agriculture Sector Optimization Model Greenhouse Gas versions for the Technical Working Group on Agricultural Greenhouse Gases (T-AGG).
- Conducted a webinar series for 33 expert scientists to determine the scientific certainty of various agriculture greenhouse gas mitigation activities. Contributed to ongoing collaborative policy and scientific analyses for T-AGG that resulted in the production of two scientific manuscripts and multiple reports for policy makers.
- Supervised research efforts of three graduate student researchers as part of the Center for Sustainability and Commerce.
- *R (ggplot2, data.table), MATLAB, SQL*

Research Assistant at the Virginia Institute of Marine Science *(Aug. 2005 to Jan. 2010)*

- Conducted chemical analysis of fish for polychlorinated biphenyls (PCBs) and other organic toxins for the Virginia Department of Environmental Quality. Delivered data to the Virginia Department of Health used to generate fish consumption warnings throughout the state.
- *MATLAB, SAS, STATA*

EDUCATION:

Master of Public Policy (2010) from the College of William and Mary

Master of Science (2010) from the College of William and Mary

Thesis: The Toxicology of Decabromodiphenyl Ether in Avian Embryos: Disposition of the Flame Retardant BDE-209 in Yolk-Injected Chicken Embryos (Gallus gallus)

- Designed and conducted a collaborative toxicology study of the brominated flame-retardant BDE-209 in partnership with the U.S. Geological Survey's Patuxent wildlife Research Center in Beltsville, MD. Delivered data to the US Environmental Protection Agency (EPA) used to support the removal of the product (BDE-209) from the market.

Bachelors of Science (2003) from the University of Massachusetts

PEER-REVIEWED JOURNAL ARTICLES

Mazotta, M.J; Wainger, L.; Sifleet, S.; Petty J.T.; Rashleigh. B. (2015). Benefit transfer with limited data: An application to recreational fishing losses from surface mining. Ecological Economics (in press) doi:10.1016/j.ecolecon.2015.09.018

Pendleton, L., Donato, D. C., Murray, B. C., Crooks, S., Jenkins, W. A., Sifleet, S. D., et al. (2012). Estimating global "blue carbon" emissions from conversion and degradation of vegetated coastal ecosystems. *PLoS ONE, 7*(9), e43542. doi:10.1371/journal.pone.0043542.

Doccola, J. J, Bristol, E. B., Sifleet, S. D., Lojko, J., & Wild, P.M. (2007). Efficacy and duration of trunk-injected imidacloprid in the management of hemlock woolly adelgid. *Arboriculture and Urban Forestry, 33*(1), 12–21.

Doccola, J. J., Ramasamy, I., Castillo, P., Taylor, C., & Sifleet, S. (2005) Efficacy of arborjet viper microinjections in the management of hemlock woolly adelgid (*Adelges tsugae*). *Journal of Arboriculture, 31,* (1), 203–206.

PRESENTATIONS AND PROCEEDINGS

Sifleet, S.D. (2016, May). *Developing Shiny Applications: Using Leaflet in R*. Presented at the Free and Open Source Software for Geopspatial (FOSS4G) Conference in Raleigh NC.

Sifleet, S.D. (2014, April). *Incorporating values and assessing social and environmental trade-offs in managing for ecosystem services.* Presented during a workshop at Estimating Recreation Demand in the EnviroAtlas, Socio-Environmental Synthesis Center (SESYNC), Annapolis, MD (remote).

Sifleet, S. D. (2013, August). *A standard assessment framework for ecosystem services.* Presented during a workshop at EnviroAtlas: Connecting Ecosystems, People, and Wellbeing, National Center for Ecological Analysis and Synthesis (NCEAS), Santa Barbara, CA.

Sifleet, S., Neale, A., Wainger, L., & Mehaffey, M. (2013, July). *Measuring ecosystem rarity to target conservation efforts.* Presented at the International Congress for Conservation Biology, Baltimore, MD.

Sifleet, S. D. S., McKernan, M., Rattner, B., Hale, R. C., LaGuardia, M., & Luellen, D. (2009, June). Presented at a platform session of the 19th Annual Meeting of Society of Environmental Toxicology and Chemistry (SETAC) Europe, Göteborg, Sweden.

Sifleet, S. D. S., McKernan, M., Rattner, B., Hale, R. C., LaGuardia, M., & Luellen D. (2008). *Comparative species analysis using PCA of biotransformation and distribution of BDE-209 in yolk-injected chickens (Gallus gallus), kestrels (Falco Sparverius), and mallards (Anas platyrhynchos)*. Poster presented at the 29th Annual North American Meeting of the Society of Environmental Toxicology and Chemistry (SETAC), Tampa, FL.

TECHNICAL REPORTS

Sifleet, S. D. (2013, September). *North American blue carbon scoping study*. Prepared for the Commission for Environmental Cooperation.

Sifleet, S. D., Pendleton, L., & Murray, B. C. (2011, May). *State of the science on coastal blue carbon: A summary for policy makers*. Prepared for Duke University's Nicholas Institute for Environmental Policy Solutions. NI R 11-06. 42 pages.

Murray, B. C., Pendleton, L., Jenkins, W. A., & Sifleet, S.D. (2011, April). *Green payments for blue carbon: Economic incentives for protecting threatened coastal habitats*. Prepared for Duke University's Nicholas Institute for Environmental Policy Solutions. NI R 11-04. 42 pages.

Eagle, A. J., & Sifleet, S. D. (2011, April). *T-AGG survey of experts: Scientific certainty associated with GHG mitigation potential of agricultural land management practices*. Prepared for Duke University's Nicholas Institute for Environmental Policy Solutions. NI R 11-05. 34 pages.