

# STACY LEE

(646) 481 8838 📞  
stacy-lee.github.io 🌐  
linkedin.com/in/lee-stacy 📄  
stacy.lee013@gmail.com ✉

Languages	Proficient: <b>Python, R, SQL</b>	Basic: <b>Scala, HTML5/CSS/JavaScript, SAS</b>
Libraries	Python: <b>NumPy, Pandas, Sklearn, SciPy, StatsModels, Matplotlib</b>	R: <b>Tidyverse, Caret</b>
Data Tools	<b>Hive, Google BigQuery, Tableau, Trifacta, Azure, AWS, Spark, UNIX/Bash, Git</b>	
Education	University of Illinois at Urbana-Champaign	
	<b>M.S. in Statistics Analytics Concentration</b>	Dec 2018
	<i>Relevant Courses: Machine Learning, Computational Statistics, Data Mining</i>	
	<b>B.S. in Civil Engineering Systems Engineering Concentration</b>	May 2017
Experience	<b>Carat, a Dentsu Company</b> New York, NY	
	<b>Analyst, Data Analytics Data Science Impact Team</b>	Sept 2019 - Present
	<ul style="list-style-type: none"><li>Lead the strategy for creating a new composite scoring metric using principal component analysis, selecting test &amp; control market pairs to measure advertising impact, &amp; creating cost estimates to generate baselines of significant sample sizes for Fortune 100 clients</li><li>Implement additive time series model using over 50 variables for causal impact analysis that received positive client feedback with approval &amp; became an official team product</li><li>Oversee end-to-end research &amp; development of XGBoost model for production forecasting</li><li>Improve online ad performance by over 20% through optimal frequency &amp; creative analyses</li><li>Develop data automation code &amp; flows to provide daily or weekly performance reporting</li><li>Create data visuals to investigate discrepancies and gain deeper insights on media trends</li><li>Collaborate with account team to ensure proper tagging taxonomy for reporting accuracy</li><li>Launch first monthly data science learning program within team to encourage growth</li></ul>	
	<b>Ameren</b> Champaign, IL	
	<b>Data Science Innovation Intern Innovationdr Team</b>	Jan 2018 - Dec 2018
	<ul style="list-style-type: none"><li>Led a 5-person team project for identifying individuals out of 1.4 million customers with high propensities of enrolling &amp; saving in energy savings programs using demographic data to create a target list of recipients who will receive the promotional bill insert<ul style="list-style-type: none"><li>Built custom tools for exploratory data analysis to gain insights &amp; formulate a data strategy</li><li>Implemented random forest for 1% imbalanced class ratio and improved recall by 90%</li><li>Applied Bayesian statistics with logistic regression for targeted marketing deliverable</li><li>Surfaced and presented insights on customer trends through storytelling to stakeholders</li></ul></li><li>Researched methods to replace the traditional utility pole health assessment process<ul style="list-style-type: none"><li>Converted &amp; reformatted raw data collected from sensors to facilitate analytics</li><li>Utilized clustering methods to find patterns in pole health based on movement</li></ul></li><li>Assisted with recommender system to estimate solar PV adoption for distribution planning</li><li>Mentored team members on machine learning, statistics, and programming in Python or R</li><li>Executed &amp; documented ETL processes in Amazon Web Services EC2 for team efficiency</li></ul>	
Competitions	<b>Synchrony Financial Datathon</b> Top 5 of 25 Teams with Best MSE Score	April 2018
	<ul style="list-style-type: none"><li>Researched the impact of federal interest rates on home improvement spending</li><li>Implemented Elastic Net with Five-Fold Cross Validation for the final submission</li></ul>	
Projects	<b>Machine Learning &amp; Computational Statistics Projects</b> [stacy-lee.github.io/ds/projects.html]	
	<ul style="list-style-type: none"><li>Monte Carlo method to identify Weibull PDF in Traffic Volume Counts 2012 NYC Open Data</li><li>Time series with fourier seasonality on Walmart sales data achieved WMAE less than 1630</li><li>Logistic regression for movie review analysis with NLP achieved AUC greater than 0.95</li></ul>	
	<b>Algorithms</b> (Written From Scratch) Random Forest, k-NN, Lasso Regression, Apriori	