

# STACY LEE

(917) 617 9342 ☎  
stacy-lee.github.io 🌐  
linkedin.com/in/lee-stacy in  
stacy.lee013@gmail.com ✉

Languages	Proficient: <b>Python, R, SQL</b>	Basic: <b>HTML5/CSS/JavaScript, C++, SAS</b>
Libraries	Python: <b>NumPy, Pandas, Sklearn, SciPy, StatsModels, Matplotlib</b>	R: <b>Tidyverse, Caret</b>
Data Tools	<b>Google Cloud Platform, Google Data Studio, Tableau, Trifacta, Azure, AWS, UNIX/Bash, Git</b>	
Marketing	<b>Google Marketing Platform, Facebook, Twitter, Adobe Analytics, Mediaocean, Search Ads 360</b>	
Experience	<b>Carat USA</b> New York, NY <b>Analyst, Data Analytics Impact Team</b> Sept 2019 - Present <ul style="list-style-type: none"><li>Build &amp; maintain data pipelines using Python &amp; SQL for marketing dashboards</li><li>Collaborate with account teams to solve data issues &amp; present reports to client with insights</li><li>Optimize Python code to be reusable for weekly or daily report automation &amp; ad-hoc requests</li><li>Monitor automated jobs in Trifacta &amp; troubleshoot issues from 5+ disparate data sources</li><li>Lead the strategy for selecting test &amp; control market pairs to measure TV advertising impact &amp; the research for creating cost estimates to generate baselines of significant sample sizes</li><li>Assisted with implementing XGBoost model to forecast production volume &amp; causal impact analysis using additive time series model to estimate marketing impact</li><li>Improve online ad performance by over 20% through optimal frequency &amp; creative analyses</li><li>Analyze web &amp; search data to understand production impact from TV &amp; radio advertising</li></ul> <b>Ameren</b> Champaign, IL <b>Data Science Innovation Intern Analytics 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 administered by the business partner using demographic dataset of mixed data types with positive feedback from partner<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 teammates on machine learning, statistics, and programming in Python or R</li><li>Executed &amp; documented ETL processes in AWS EC2 for efficiency</li></ul>	
Education	University of Illinois at Urbana-Champaign <b>M.S. in Statistics Analytics Concentration</b> Dec 2018 <b>B.S. in Civil Engineering Systems Engineering Concentration</b> May 2017	
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>	

**Algorithms** (Written From Scratch) Random Forest, k-NN, Lasso Regression, Apriori