

STACY LEE

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

Languages	Proficient: Python, R, SQL	Basic: SAS, HTML5/CSS/JavaScript, C++
Libraries	Python: NumPy, Pandas, Sklearn, SciPy, StatsModels, Matplotlib	R: Tidyverse, Caret
Tools	GCP, Azure, Hive, Spark, Tableau, AWS EC2, UNIX/Bash, Git	
Experience	Carat USA New York, NY Analyst, Data Analytics <i>Impact Team</i> Sept 2019 - Present <ul style="list-style-type: none">Lead the research for creating cost estimates to attain significant sample sizes for generating baselines in marketing campaign planningImprove online ad performance over 20% through optimal frequency & creative analysesCollaborate with cross-functional teams to present media performance reports with insightsImplement time series models for forecasting production volume & causal impact analysis for estimating marketing impact by including external factors from researchOptimize Python code to be reusable for WoW or daily report automation and ad-hoc requestsBuild & maintain data pipeline using Python and SQL for media performance dashboardsDesign strategy for test & control market pairing to measure marketing ROIAnalyze web data in Adobe Analytics & Search Ads 360 for marketing attribution Ameren Champaign, IL Data Science Innovation Intern <i>Analytics Team</i> Jan 2018 - Dec 2018 <ul style="list-style-type: none">Led a 5-person team project for identifying individuals out of 1.4 million customers with high propensities of enrolling & 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">Built custom tools for exploratory data analysis to gain insights & formulate a data strategyImplemented random forest for 1% imbalanced class ratio and improved recall by 90%Applied Bayesian statistics with logistic regression for targeted marketing deliverableSurfaced and presented insights on customer trends through storytelling to stakeholdersResearched methods to replace the traditional utility pole health assessment process<ul style="list-style-type: none">Converted & reformatted raw data collected from sensors to facilitate analyticsUtilized clustering methods to find patterns in pole health based on movementAssisted with recommender system to estimate solar PV adoption for distribution planningMentored teammates on machine learning, statistics, and programming in Python or RExecuted & documented ETL processes in AWS EC2 for efficiency	
Education	University of Illinois at Urbana-Champaign M.S. in Statistics <i>Analytics Concentration</i> Dec 2018 B.S. in Civil Engineering <i>Systems Engineering Concentration</i> May 2017 <i>Minor in Mathematical Statistics</i>	
Competitions	Synchrony Financial Datathon Top 5 of 25 Teams with Best MSE Score April 2018 <ul style="list-style-type: none">Researched the impact of federal interest rates on home improvement spendingImplemented Elastic Net with Five-Fold Cross Validation for the final submission	
Projects	Machine Learning & Computational Statistics Projects [stacy-lee.github.io/ds/projects.html] <ul style="list-style-type: none">Monte Carlo method to identify Weibull PDF in Traffic Volume Counts 2012 NYC Open DataTime series with fourier seasonality on Walmart sales data achieved WMAE less than 1630Logistic regression for movie review analysis with NLP achieved AUC greater than 0.95 Algorithms (Written From Scratch) Random Forest, k-NN, Lasso Regression, Apriori	