

STACY LEE

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

Languages	Proficient: Python, R, SQL	Basic: Scala, HTML/CSS/JavaScript, C++
Data Tools	Spark, Tableau, Superset, Google BigQuery, AWS, Airflow, Jenkins, Hive, Azure, UNIX, Gitlab	
Experience	Nielsen New York, NY Senior Data Scientist Nielsen Marketing Cloud Team June 2021 - Present <ul style="list-style-type: none">Lead the model evaluation methodology of Nielsen's identity graph for audience segmentsBuild varying model iterations of the identity graph for research & experimentationDeploy automated dashboards for monitoring segmentation model performanceImplement core metrics & identify areas for monitoring discrepancies in data pipelineCollaborate cross-functionally with product & operations teams to align on implementation	
	Dentsu New York, NY Data Scientist Marketing Effectiveness Team Sept 2019 - June 2021 <ul style="list-style-type: none">Led the strategy for creating a new composite scoring metric using principal component analysis, selecting test & control market pairs to measure advertising impact, & creating cost estimates to generate baselines of significant sample sizes for Fortune 100 clientsBuilt additive time series model using over 50 variables for causal impact analysis to estimate incremental lift with results that received positive client feedbackOversaw end-to-end research & development of XGBoost model for production forecastingImproved online ad performance by over 20% through optimal frequency & A/B testingDeveloped automation code & data flows using APIs for presenting daily or weekly reportsCreated data visuals to investigate discrepancies and gain deeper insights on media trendsCollaborated across teams to ensure proper tag taxonomy for Facebook & Twitter ads	
	Ameren Champaign, IL Data Science Innovation Intern Innovation Team 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 using demographic data of 150+ features to create a target list of recipients who will receive the promotional bill insert<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 deliverablesResearched methods to replace the traditional utility pole health assessment process<ul style="list-style-type: none">Utilized clustering methods to find patterns in pole health based on movementAssisted with recommender system to estimate solar PV adoption for distribution planningMentored team members on machine learning, statistics, and programming in Python or RCreated & documented code to extract large datasets from AWS S3 for team efficiency	
Education	University of Illinois at Urbana-Champaign M.S. in Statistics Data Science & Analytics Concentration Dec 2018 Relevant Courses: Machine Learning, Computational Statistics, Data Mining B.S. in Civil Engineering Systems Engineering Concentration May 2017	
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	