

STACY LEE

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

Summary	Data science and analytics professional with 4+ years of experience applying data science methods to drive decision-making and building data pipelines for product & marketing analytics.		
Languages	Proficient: Python, R, SQL	Basic: Scala, HTML/CSS/JavaScript, C++	
Data Tools	Tableau, Power BI, AWS, Spark, Airflow, BigQuery, Hive, Azure, dbt, Jenkins, Git		
Education	University of Illinois at Urbana-Champaign M.S. in Statistics <i>Data Science & Analytics Concentration</i> <i>Relevant Courses: Machine Learning, Computational Statistics, Data Mining</i> B.S. in Civil & Environmental Engineering <i>Transportation & Systems Concentration</i>		
Experience	Nielsen	New York, NY	
	Senior Data Science Analyst <i>Nielsen Marketing Cloud Team</i>	June 2021 - Jan 2023	
	<ul style="list-style-type: none">○ Led the model evaluation & deployment of graph product features built on terabytes of data○ Created & evaluated model iterations of graph product produced by clustering algorithm○ Built automated dashboards to monitor audience segment models & gain weekly insights○ Scoped & planned KPI metrics for monitoring anomalies and discrepancies in data pipeline○ Debugged issues in modeling pipeline that removed data leakage & reduced query runtime○ Presented & documented model evaluation methodology & results for reproducibility○ Mentored & reviewed the Python code & queries of less experienced colleagues		
	Dentsu	New York, NY	
	Data Science Analyst <i>Marketing Effectiveness Team</i>	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 project sample size baselines for Fortune 100 financial & retail campaigns○ Built additive time series model using over 50 variables for causal impact analysis to estimate incremental lift from media campaigns with results that received positive client feedback○ Implemented feature engineering in predictive models improving accuracy over 10%○ Oversaw end-to-end research & development of XGBoost model for marketing mix modeling○ Improved online ad performance by over 20% through optimal frequency & A/B testing○ Automated ETL data pipeline from multiple sources using APIs for daily or weekly reports○ Created data visuals to investigate discrepancies & gain insights for storytelling○ Collaborated across teams to ensure proper website tags for tracking in Adobe Analytics		
	Ameren	Champaign, IL	
	Data Science Innovation Intern <i>Innovation 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 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">- Applied Bayesian statistics with logistic regression for targeted marketing deliverables- Implemented random forest for 1% imbalanced class ratio & improved recall over 50%- Created custom tools for exploratory data analysis to gain insights & form a data strategy○ Utilized clustering methods to find patterns in pole health based on movement from sensors		
	Projects	Machine Learning & Computational Statistics Projects [stacy-lee.github.io/ds/projects.html] <ul style="list-style-type: none">○ Logistic regression for movie review analysis with NLP achieved AUC greater than 0.95○ Monte Carlo simulation to identify Weibull PDF in Traffic Volume Counts 2012 NYC Open Data Algorithms (Written From Scratch) Random Forest, k-NN, Lasso Regression, Apriori	