Da Som Lee

188 St. Nicholas Ave #34, New York, NY, 10026 · 929-225-6166 · dl3153@columbia.edu · https://dldasomlee.github.io

EDUCATION

Columbia University, Graduate School of Arts and Sciences　　　　　　　New York, NY
M.A. in Data Science　　　　　　　　　　　　　　　　　　　　　　　　*Expected Dec 2018*
- *Relevant coursework*: Algorithms for Data Science, Probability and Statistics for Data Science, Modern Data Structure

Korea University Institute for Continuing Education in Mathematics　　　Seoul, South Korea
- *Relevant coursework:* Linear Algebra, Probability and Statistics　　　　　　*June – Aug 2016*

NYU Steinhardt School of Culture, Education, and Human Development　　New York, NY
B.S. in Media, Culture, and Communication　　　　　　　　　　　　　　*Dec 2013*

WORK EXPERIENCE

Google　　　　　　　　　　　　　　　　　　　　　　　　　　　　Seoul, South Korea
Account Coordinator　　　　　　　　　　　　　　　　　　　　　　*Aug 2015 – Oct 2016*
- Generated $3M increase in revenue through collaboration with industry head and data scientist to build a predictive model using internal search tool and YouTube data set to create competitive brand positioning for major advertisers
- Investigated the effect of shifting ad spending from TV to YouTube on audience reach via meta-analysis with data analysts and drafted proposals for key clients to allocate a $10M financial plan onto digital advertising
- Examined 100K of consumer web behavior conversion paths with software engineers using Google Analytics API in **R** and increased sales on display ads by 40% through identifying opportunities to improve online lead generation
- Used **SQL** and **Python** to clean top 10 revenue generating advertisers' AdWords 100K campaign data and applied linear regression models to identify key performing indicators of banner ads performance
- Advised on 150 technical solutions weekly across all products to media agencies resulting in 95% client satisfaction

Research Assistant　　　　　　　　　　　　　　　　　　　　　　Seoul, South Korea
Department of Mathematics, Korea University　　　　　　　　　　　　　*Aug – Oct 2016*
- Assisted professor in research on financial product pricing model in a competitive financial market by gathering and interpreting the data of price elasticity and production cost; presented development progress weekly to the team

INDEPENDENT PROJECTS

Sentiment Analysis of MTA's Real-time Tweets on a Thanksgiving Holiday *(R, JSON, AWS)*
- Collected 10K data size of MTA's tweets and riders' retweets using the Twitter streaming API and Amazon's EC2
- Created plots with ggplots2 to show the frequency of tweets and found a strong association for independent and dependent variables by running a chi-square test
- Performed sentiment analysis using the NRC lexicon and measured the number of positive or negative sentiment, sum and maximum of each score, and normalized average sentiment strength.

Revenue Collection Regimes in the Extractive Industries Transparency Initiative (EITI) *(R, Python, Tableau)*
- Gathered and analyzed the EITI payment and secondary datasets of 10K to explore the pattern of effective revenue collection strategy to changes in commodity markets for 25 countries
- Performed regression analysis to run different hypothesis testing and generated multi-graphs in Tableau

Predicting User Preference for Movies *(Python)*
- Analyzed 1K data set consisting of movie reviews; built an analytical model to estimate users' movie preferences by implementing the Naïve-Bayes algorithm and achieved 5-fold cross-validation accuracy of 95%

SKILLS & CERTIFICATIONS

Skills: Python (Scikit-learn, NumPy, SciPy, Pandas), R, Java, SQL, Tableau, AdWords, YouTube and Google Analytics, MS Office
Languages: Fluent in Korean, intermediate in Chinese and Japanese
Certifications: UC Berkeley Extension: Java; Stanford University: Machine Learning; UIUC: Text Mining and Analytics; Multivariable Calculus; Tsinghua University: Combinatorial Mathematics and four other courses