Kessie Zhang

Data Scientist | Analyst
San Francisco, CA

CONTACT INFO

(415)-706-3878
kessiezhang0620@gmail.com
LinkedIn: kessie-zhang
GitHub: kessiezhang
http://kessiezhang.github.io/

SKILLS

Programming

Python, R
HTML, CSS, Javascript

Statistics

Hypothesis Testing
A/B Testing
Multivariate Testing

Data Science

Regression
Classification
Clustering
NLP
Time Series
Web Scraping

Data Visualization

matplotlib, seaborn
Plotly
Tableau
Power BI

Tools

PostgreSQL, MongoDB
MS Excel, Access
AWS, Spark
Hadoop
Mixpanel

Languages

Mandarin
Cantonese

EDUCATION

M.S. Business Analytics *June 2020*
California State University, East Bay

B.S. Applied Mathematics *August 2017*
University of California, San Diego

EXPERIENCE

Data Scientist *March 2020 - Present*
CrowdDoing | San Francisco, CA

▷ Extract products that contain ingredients with medicinal properties and their reviews from Amazon.com through web scraping and conduct data preparation
▷ Apply NLP and unsupervised learning techniques to 5+ million Amazon products reviews to evaluate products' effectiveness in stress, sleep and anxiety disorders and uncover customers purchase preferences
▷ Collaborate with data scientists and engineers to develop a content-based filtering recommender to recommend medicinal foods that users might like
▷ Interpret and discuss the findings weekly with the managers and colleagues

Data Science Fellow *July 2019 - Sept 2019*
Metis | San Francisco, CA

▷ Collaborated with other data scientists to conduct an exploratory data analysis using python and presented recommendations to identify optimal timings and street teams' allocation to maximize ticket purchases for an event
▷ Applied NLP and machine learning techniques on hotels reviews and deployed a Tableau dashboard to make booking/comparing hotels more convenient
▷ Produced a perfume recommendation system by content-based filtering and topic modeling on perfume descriptions to help customers identify similar alternatives to their favorite scents
▷ Implemented and evaluated various predictive models (tree classifiers, support vector machines, logistic regression, xgboost) on recall to determine whether or not a user will book a room
▷ Built a linear regression model on 100+ court NBA performance metrics and salaries scraped from several websites to help managers draft players with high performance-to-cost ratios

Data Analyst Intern *June 2018 - Aug 2018*
Shanghai Pudong Development Bank | Guangzhou, China

▷ Analyzed businesses and individuals' financial data to determine the level of risk associated with extending credit
▷ Built Tableau dashboard to provide insights for assessing credit risks and identify opportunities for improvement
▷ Validated the accuracy and legitimacy of clients' financial documents which reduced friction between the bank and the clients

Junior Analyst *Dec 2017 - April 2018*
Exela Technologies | San Francisco, CA

▷ Developed data analysis models using Excel to reduce operational costs and maximize profit yields
▷ Improved processing efficiency of inventory and pricing data updates on MS Access database via SQL
▷ Ingested data from different sources, and created reports using Pivot Table and Vlookup in Excel