




JAMES HU

San Francisco, CA
(386)-405-4045
contact@LJamesHu.com

 <http://LJamesHu.com/>
 <https://github.com/ljameshu>
 <https://linkedin.com/in/ljameshu/>

EMPLOYMENT

SoFi **San Francisco, California**
Data Scientist, Member **2019 - Present**

- Created core membership data structure and dashboards to track product ownership, attributes, and activity in time series for over 2 million users to understand users across 9 different financial products.
- Established data-driven customer experience system using product NPS surveys via Qualtrics, setting up data pipelines (export/ingestion), with time series, confidence intervals and NLP analysis through Tableau.
- Worked with business and engineering stakeholders to create centralized KPI infrastructure and dashboards to track hundreds of metrics for financial status, business health, and engineering progress.
- Developed core user level tables and methodologies to track user activity, retention, and behaviors and leverage that to apply machine learning to understand and improve those metrics.
- Stood up and maintained automation infrastructure and systems through Airflow and AWS.

Product Madness **San Francisco, California**
Data Scientist **2017 - 2018**

- Redesigned Tableau infrastructure and ETL processes for user acquisition team to understand and optimize core performance metrics on multi-million dollar monthly user acquisition spend.
- Created predictive revenue model to predict lifetime value with machine learning techniques to optimize user acquisition campaigns and provide signaling to intelligently adjust bids.

Independent Data Science Consulting **San Francisco, California**
Data Science Consultant **2017**

- Developed a corporate rental pricing model using Airbnb, Trulia, and different neighborhood metric data to improve pricing and understand inputs in corporate housing market.
- Applied natural language processing and clustering techniques to analyze live chat logs for a large nationwide telecommunications company on chat automation efficacy and secure a long term contract.

Galvanize **San Francisco, California**
Data Scientist in Residence **2016**

- Supported and mentored 30+ students to develop core data science, programming, and statistical skills.

Strategic Insight **New York, NY**
U.S. Data Research Associate **2014-2015**

EDUCATION

New York University - Leonard N. Stern School of Business **New York, NY**
Bachelor of Science in Finance and Economics; Minor in Mathematics **2010-2014**

PROJECTS

Heart-icle

- Used semi-supervised learning with Medium's user generated highlights to create better summaries with Gensim's Doc2Vec and Medium highlight data to capture the relationship between an article and highlight.
- Scraped over 30,000 articles from Medium and The Atlantic using BeautifulSoup and Selenium.

SKILLS

Programming Skills: Python (pandas, NumPy, scikit-learn, SciPy, BeautifulSoup, Gensim, Flask)
git (GitHub, GitLab), Selenium, Java (Intermediate)

Databases: Snowflake, Athena (Presto), Amazon Redshift, PostgreSQL, MySQL, MongoDB

Other Technical Skills: Apache Airflow, Tableau, Jupyter, Amplitude, Redash, AWS (S3, EC2, Sagemaker), Server Admin, Linux (Ubuntu, CentOS 7), Jekyll, Nginx, Wordpress, LaTeX