

Kailei Lin

linkailei655@gmail.com | +1 (323) 351-3716 | Los Angeles, CA | [Linkedin](#) | [Portfolio](#)

Technical Skills

Languages & Platforms: Python(OOP,Pandas), SQL(Snowflake,MySQL,S3, Redshift), R, Tableau, Hadoop, EC2

Statistics & Data Science: Classification, Regression, Time-Series(ARIMA), Hypothesis Testing(t-test, z-test)

Certification: AWS Cloud Practitioner

EXPERIENCE

Beacon Economics, LLC

Sep 2021 – Present

Business Data Analyst

Los Angeles, CA

- Enabled decision-making by developing a clustering model to segment and benchmark 15 economic performance metrics among 19,500 cities. Suggested change in strategic priorities, potentially boosting revenue by 18%.
- Improved profit margins through collaborative efforts with business stakeholders, conducting in-depth annual and monthly expenses analysis, scrutinizing P&L statements and generating reports using econometric models.
- Facilitated weekly business review meetings with stakeholders and product managers by presenting quantified discrepancies between target and current performance and proposed target timeline based on CAGR and regression model.
- Utilized **hypothesis testing** and **linear regression** to identify causal factors impacting business functions, suggesting operational strategies with the potential to achieve an 24% reduction in cost.
- Automated data cleaning, extraction, econometrics modeling, and visualization in Dashboards using Python and SQL.
- Mined open-source data with **Python** (20 years of 41,683 zip code level business activities data) and built a high-quality infrastructure in **S3 database**, increasing analysis credibility by 30%.
- Collaborated with Data Engineers to complete cloud transformation (Fileshare to **S3-Redshift**) and liaised between the VP and engineers to optimize cost structures, saving thousands of dollars annually.

Tusimple, INC(NASDAQ: TSP)

July 2020 – Nov 2020

Data Analyst Intern - Maps

San Diego, CA

- Designed and maintained ETL pipelines and tables that provided insights on **+160k log** Data.
- Built an automated Tableau dashboard consisting of 40 key metrics to capture product dynamics.
- Collaborated closely with product managers and algorithm team to perform root-cause analysis using **K-Nearest Neighbors** on over 1M false positives log data gathered through SQL, leading to a 15% improvement in accuracy rate.

Selected Projects

Stock Performance & Selection | Python, Time-Series Analysis

Jan 2023 – May 2023

- Constructed a multi-factor stock selection model by leveraging gathered data from Yahoo Finance and optimized the selection strategy using **SVM** and **Logistic Regression**, achieving an AUC of 0.92.

Cloud-based University Search Engine | Python, JavaScript, npm/Node.js, EC2

Jan 2020 – May 2020

- Crafted back-end database(**EC2**) and implemented **inverted indexing** schema to facilitate real-time data retrieval.
- Built a search engine UI in **jQuery**, allowing viewers to search, filter, and continuously navigate from returned results, resulting in a weekly traffic of 500 visitors and a 20% increase in email subscriptions.

EDUCATION

University of Southern California

Aug 2018 – Dec 2020

Master of Science; Major in Spatial Data Science

Los Angeles, CA

Miami University-Oxford

Aug 2014 – May 2018

Bachelor of Science; Major in Information Systems

Oxford, OH