Shiyi "Ashley" Yue

(315) 746-9070 | ashley.yue117@gmail.com | CA (open to relocation) | Portfolio | LinkedIn | GitHub

PROFILE

MS in Business Analytics '23 from UC Davis | Proficient in SQL, Python, Excel, R, and Tableau | Data-driven problem solver | Results-oriented, attention to detail, team player

Skills: Data Wrangling, Statistical Modeling, Hypothesis Testing, EDA, ETL, Machine Learning, Data Visualization, A/B Testing, Time Series Analysis, Image Classification, Web Scraping

Technologies: Advanced SQL, Python (Numpy, Pandas, Seaborn, Scikit-Learn, TensorFlow, PySpark), R, Tableau, Advanced Excel, NoSQL, Jupyter, AWS, GCP, GitHub, HTML, MATLAB, SPSS, Stata, G Suite, MS Office

Certificates: Tableau Desktop Specialist, SQL(Advanced), Machine Learning Specialization, AWS Cloud Practitioner

EDUCATION

University of California - Davis — **San Francisco, CA**
Master of Science, Business Analytics (3.70/4.0) — Aug. 2022 - Jun. 2023
Highlighted Coursework: Data Management, Machine Learning, Advanced Statistics, Data Visualization, Big Data

Southwest University of Political Science and Law — **Chongqing, CHN**
Bachelor of Economics, Economic Statistics (3.77/4.0) — Sept. 2018 - Jun. 2022
Highlighted Coursework: Statistics, Data Mining, Time Series Analysis, Econometrics, Finance, Accounting

State University of New York at Oswego — **Oswego, NY**
Exchange Program, Psychology & Communication (3.83/4.0) — Aug. 2021 - May 2022

PROFESSIONAL EXPERIENCE

Fashom — **Miami, FL**
Data Scientist (Practicum Project) — Sept. 2022 – Jun.2023
Leveraged AI and data analysis to help an online retail startup transfer to a B2B SaaS that offers API solutions

- Led a team in boosting clothing image classification accuracy by 40% with Python TensorFlow and Google CoLab, leveraging data augmentation, GPU cloud computing, and ResNetV2 transfer learning
- Developed a personalized recommendation engine featuring Market Basket Analysis using Python and SQL
- Performed data validation and EDA on 63k+ clothing inventory data from AWS Workspace, using SQL and Python
- Collected and managed 80K+ clothing image data from multiple sources using Python and Chrome extensions
- Effectively communicated insights to technical and non-technical stakeholders using Tableau and Google Slides

PricewaterhouseCoopers (PwC) — **Shenzhen, CHN**
Business Analyst Intern — Sept. 2021 – Dec. 2021
Satisfied an E-commerce client with data modeling and promotion strategy design

- Applied K-means Clustering in Python to identify customer segments for tailored promotion policy design
- Designed an A/B test aimed at new users to identify promotion policies optimizing user conversion rates
- Employed SQL to compute conversion/retention rates and generate summary statistics for demographic variables
- Utilized Python and SQL to perform EDA on 210k+ daily user engagement data, extracting key business insights

Founder Securities — **Shenzhen, CHN**
Data Analyst Intern — Feb. 2021 – Jul. 2021
Contributed to data analysis and data modeling projects for FMCG clients

- Designed ETL processes with SQL and Python scripts based on business requirements, ensuring data accuracy
- Built an ARIMA model in R on 327K+ sales records for inventory prediction, achieving an 84% test set accuracy
- Collaborated with clients' data team to ensure seamless data integration and alignment with their business objectives
- Developed interactive Tableau dashboards to visualize KPIs and sales performance for stakeholders

Fanhua Finance Insurance — **Chengdu, CHN**
Marketing Analyst Intern — Oct. 2020 – Feb. 2021
Improved marketing strategies utilizing data analysis at the ads team

- Monitored ad data and analyzed target customer profiles across channels using Excel pivot tables and charts
- Improved campaign effectiveness with competitive analysis and trends identification, achieving a 15% CTR boost
- Optimized ads budget allocation using Excel optimization solver, resulting in a 10% increase in total ROI

PROJECTS

Customer Repurchase Analysis: Developed Random Forest models to identify customer repurchase propensity scores
Trending Music Analysis: Leveraged Logistic Regression and Confusion Matrix to predict and evaluate trending classes
Top Pizzeria in SF: Utilized Python BeautifulSoup, MongoDB, and API to request, transform, and store pizzeria data

Find more projects on My Portfolio↗