

Jiayi Zhu

jyzyo04@gmail.com | Seattle, WA | github.com/svedison | linkedin.com/in/jiayi-zhu-549884298

PROFESSIONAL SUMMARY

UW senior in Applied & Computational Mathematical Sciences (Discrete Math & Algorithms), graduating June 2027 and available to start Summer 2027, with three software engineering internships centered on data platforms and pipelines. Independently designed an end-to-end master-data-management pipeline on Amazon Aurora PostgreSQL with golden records and field-level lineage, engineered a distributed Apache Spark pipeline on Databricks processing 4.5M+ records, and built Python ETL pipelines with API ingestion and data validation integrating 10+ external datasets. Strong CS fundamentals (data structures & algorithms, probability, statistical modeling) with substantial SQL, Python, and big-data-engine experience from internships, projects, and hackathons.

EDUCATION

University of Washington

B.S. in Applied & Computational Mathematical Sciences

Seattle, WA

Expected June 2027

Concentration: Discrete Math & Algorithms · GPA: 3.70/4.00 · Dean's List

Coursework: Data Structures & Algorithms, Probability, Statistical Modeling, Machine Learning.

Leadership: Student Leader, Chinese Student Organization; led cross-functional marketing initiatives reaching 2,000+ students.

WORK EXPERIENCE

Symetra

Bellevue, WA

Data Engineer Intern

Jul 2026 – Present (expected Oct 2026)

- Independently designed, built, and optimized an end-to-end master data management (MDM) pipeline on Amazon Aurora PostgreSQL; significantly reduced runtime with batch processing and validated the proof of concept on hundreds of thousands of production-database records.
- Implemented deterministic/probabilistic matching, Union-Find clustering, and attribute-level survivorship to reconcile three legacy insurance-agent sources into golden records with field-level lineage.
- Developed a full-stack data stewardship application using FastAPI and React/TypeScript, enabling field overrides, record splitting, audit trails, and match tuning; demonstrated to senior leadership.
- Validated application functionality with 134+ unit tests and built a synthetic-data generator with invariant validation to support testing.

Capgemini

Bellevue, WA

Software Engineer Intern

Jun 2026 – Jul 2026

- Led PostgreSQL database development for a multi-agent automotive supply-chain automation prototype in the AI Center of Excellence; the team demonstrated the working prototype to senior leadership.
- Developed FastAPI services and synthetic-data ingestion components to simulate supply-chain behavior and support future production integration.
- Built a full-stack security-vulnerability automation proof of concept for a major retail client with FastAPI and React, demonstrated to senior leadership for potential production adoption.

Eye Spy Foundation

Jacksonville, FL

Software Engineer Intern

Jun 2025 – Aug 2025

- Built Python pipelines with REST API ingestion and data validation to integrate 10+ external datasets and 50K+ transactional records; reduced batch preprocessing time 30%, from 5.0 to 3.5 hours.
- Developed a Python/SQL machine-learning scoring system to segment campaign opportunities and automate performance evaluation, improving targeting precision 35%.

PROJECTS

Climate Yield Anomaly Detection – Databricks ML Hackathon

Seattle, WA

Data Science & ML Contributor

Feb 2026

- Engineered a distributed Python/Apache Spark pipeline on Databricks processing 4.5M+ USDA and NOAA records for crop-yield modeling and anomaly detection.
- Trained XGBoost regression models on weather variables and applied SHAP to explain feature contributions to non-technical audiences.

Creme Pudding Pet Shop – E-commerce Startup

Seattle, WA

Founder & Operator

Mar 2025 – Present

- Ran pricing A/B tests using formal statistical hypothesis testing, increasing click-through rate 28% (2.5% to 3.2%); built analytics dashboards tracking acquisition cost, conversion, and demand trends.

TECHNICAL SKILLS

Languages: Python, Java, TypeScript, JavaScript, SQL

Frameworks & APIs: FastAPI, React, REST APIs

Cloud & Data: AWS (Amazon Aurora, S3), PostgreSQL, Apache Spark, Databricks, ETL

Machine Learning: XGBoost, SHAP, Pandas, NumPy

Developer Tools: Git, Azure DevOps, Docker, Linux/Unix, unit testing, debugging, AI-assisted development