

SKILLS

LANGUAGES

Python

Scala

Java

TypeScript

SQL

C#

BIG DATA

Spark

PySpark

Apache Airflow

ETL

Data Pipelines

CLOUD & AWS

AWS Lambda

S3

Redshift

API Gateway

DATABASES

NoSQL

Data Modeling

Data Warehousing

Data Quality

Data Migration

AI / AGENTS

Amazon Bedrock

Strands

Agentic AI

LLM Integration

OTHER

ML Engineering

Data Analysis

Mentorship

EDUCATION

University of La Verne

Master's degree, Information Technology

2013 – 2015

Guangdong University of Technology

Bachelor's degree, Business Administration & Management

2006 – 2010

Udacity

Nano Degree — Machine Learning Engineer

May 2021 – Aug 2021 · Passed

Udacity

React & Web Development

Oct 2023 · Pass

SUMMARY

Senior Data Engineer with 8+ years of experience designing and scaling cloud-based big data infrastructure, ETL pipelines, and distributed data warehouses. Currently technical lead for data infrastructure within **Amazon's FinTech organization**, where I architected a **multi-account, multi-region framework spanning 11 AWS accounts** and a trusted **cost-of-goods-sold platform delivering financial visibility to 600+ programs**. Previously led architecture of marketing measurement ML pipelines achieving a **90% reduction in delivery latency**, tuned Spark workloads generating **millions of dollars in annual savings**, and built agentic AI applications with **Amazon Bedrock and Strands**. Recognized as a **top-tier performer in 2 consecutive Amazon annual reviews (2024 & 2026)** across distinct internal teams. Proven mentor across cross-functional teams of data scientists, economists, and engineers. Deep expertise in Python, Spark, AWS, and large-scale data processing.

EXPERIENCE

Amazon

Full-time · 4 yrs 5 mos

Sep 2025 – Present

Senior Data Engineer — FinTech

Seattle, WA · On-site

- Technical lead for strategically critical data infrastructure initiatives within Amazon's **FinTech** organization, driving platform modernization and financial data visibility at scale.
- Designed a **multi-account, multi-region framework** that standardized data movement and quality validation across **9 relational database clusters, 2 data warehouse clusters, and 11 AWS accounts** — establishing a consistent, auditable foundation for cross-organizational data migration.
- Led the architecture of a trusted data platform providing visibility into **cost-of-goods-sold (COGS) data across 7 product lines**, enabling financial insights for **600+ programs**.

Python · PySpark · AWS · Data Quality · Data Migration · Multi-Account Architecture

Data Engineer II — Alexa Device

Nov 2024 – Aug 2025

★ Top Tier Performance — Amazon Annual Review 2026

Seattle, WA · On-site

- Built agentic AI applications using **Amazon Bedrock** and **Strands framework**, powering intelligent multi-step automation for internal customer-facing products.
- Led design and development of a scalable cloud-based big data processing infrastructure for the Alexa Device team, optimizing data processing performance and scalability.
- Directed the design of event-driven data processing workflows, ensuring real-time and reliable data handling.
- Enhanced and automated legacy workflows, eliminating manual intervention and improving operational efficiency.

Python · TypeScript · PySpark · AWS · Amazon Bedrock · Strands

Software Engineer — Marketing Measurement

Sep 2023 – Oct 2024

Irvine, CA · Hybrid

- Designed and developed new features for the Marketing Measurement Client Portal, improving functionality and user experience.
- Developed and maintained API Gateway endpoints to enable seamless, secure data access.
- Created automation processes for one-click measurement solutions, streamlining workflows and reducing manual effort.
- Refactored unit tests for production packages, improving robustness and reliability.

Java · API Gateway · Software Development

Data Engineer II — Marketing Measurement

Jan 2022 – Sep 2023

★ Top Tier Performance — Amazon Annual Review 2024

Irvine, CA · Remote

- Led design and development of data pipelines for a marketing measurement ML model, achieving a **90% reduction in delivery latency**.
- Optimized slow-running Spark jobs, resulting in **millions of dollars in annual savings** and significantly reduced software delivery latency.
- Designed and developed automation solutions for data warehousing ETL workflows, increasing efficiency and reliability.
- Provided data engineering support for data scientists and economists, enhancing their analytical capabilities.
- Mentored junior engineers and interns, fostering technical growth across the team.

Python · AWS Lambda · Spark · Redshift · ETL · Data Warehousing

American Honda Motor Company

Contract · 3 yrs 1 mo · Torrance, CA

Senior Data / Database Engineer

Sep 2019 – Jan 2022

- Acted as lead database engineer, supporting and developing back-end databases for the consumer application team.
- Led data warehouse engineering efforts, delivering business intelligence solutions for the analytics team.
- Developed and maintained high-quality data pipelines and ETL processes across multiple data sources.

Python · C# · Data Modeling · ETL · BI

Database Engineer

Jan 2019 – Sep 2019

- Collaborated with project teams on database design and program development.
- Provided BI solutions by creating and maintaining reports and data visualizations.
- Debugged and supported existing ETL processes and database programs.

C# · SQL · Data Modeling

HELM360

Full-time · Aug 2018 – Jan 2019

Database Engineer

Aug 2018 – Jan 2019

- Designed and maintained relational databases and data models supporting business applications.

C# · SQL · Data Modeling