

Erica Vaughn

Data Scientist | Cloud Solutions Architect | Scalable AI Systems

Chicago, IL

✉ Ericat649@gmail.com | [linkedin.com/in/erica-vaughn-9b820a13b](https://www.linkedin.com/in/erica-vaughn-9b820a13b)

Professional Summary

Cloud-native Data Scientist and AI Consultant with a strong foundation in architecting production-scale ML systems across Azure and Databricks. Experienced in building CI/CD pipelines, real-time and batch ML workflows, and enterprise-level data platforms that transform operations. Proven track record delivering results in retail environments—reducing costs, optimizing inventory, and scaling analytics adoption. Excels at bridging engineering, analytics, and business stakeholders to deliver actionable insights.

Technical Skills

Languages & Modeling: Python, R, SQL, ARIMA, LSTM, XGBoost, TensorFlow, PyTorch, NLP, Causal Inference

Cloud & Data Platforms: Azure (ML, Data Factory, Databricks, DevOps), AWS, Hadoop, Spark

MLOps & Engineering: CI/CD Pipelines, Azure DevOps, Docker, Kubernetes, Feature Engineering, DataMart Design

Visualization & Tools: Power BI, DAX, PowerApps, Power Pages, A/B Testing, Experiment Design

Professional Experience

Data Scientist | Cloud + AI Consultant

NUI Ventures – Remote | Dec 2024 – Present

- Ingested and transformed data through a full medallion architecture (Bronze → Silver → Gold) in Databricks, enabling clean, structured, and analytics-ready datasets for enterprise consumption.
- Developed a scalable Data Integrity Framework using PySpark to automate data quality checks across multiple layers, integrated with Databricks workflows for daily monitoring and alerting.

- Built and published stakeholder-driven Power BI reports using Gold layer data from Databricks, with iterative development based on business feedback.
- Cataloged and enriched metadata for MDM tables using Unity Catalog, supporting Generative AI use cases by improving discoverability and context for enterprise data assets.
- Deployed real-time data quality alerting for key metrics such as null value thresholds, primary key violations, schema drift, and record count mismatches, with automated email notifications.
- Partnered with data analysts to refine and stabilize semantic models, ensuring reporting continuity and alignment after upstream schema changes.

Data Scientist, Advanced Analytics

Readerlink Distribution Services – Oak Brook, IL | Dec 2022 – Present

- Led the development of a next-gen DataMart, consolidating legacy systems and modern ERP data into a unified, scalable operational platform
- Utilized SQL and Power BI to deliver actionable insights to leadership, driving high-impact decision-making across demand planning and logistics
- Developed and deployed advanced forecasting models (ARIMA, LSTM, XGBoost) for 180+ SKUs using Azure ML and Databricks, increasing forecast accuracy by 15%
- Created a self-service AI demand forecasting portal (PowerApps + Power Pages), reducing manual work by 40% and enabling real-time scenario planning
- Automated AR/AP reconciliation using ML pipelines, contributing to \$3M in annual savings
- Implemented secure role-based access and CI/CD workflows for governed Power BI delivery
- Collaborated cross-functionally with planning, sales, and tech teams to increase inventory turns and reduce stockouts

Key Projects

AI Demand Planner: Designed a forecasting tool integrated into Power BI that optimized inventory distribution across 3 major retail outlets.

Automated AI DataMart: Created a natural language interface enabling SQL generation from user queries in Azure Databricks.

Education

M.S. in Data Science – Data Engineering (Expected Aug 2025)
Western Governors University

MBA (2023) | B.S. in Mathematics Education (2021)
Western Governors University

Certifications

Google Data Analytics Professional Certificate – Coursera, Aug 2022