

## PROFESSIONAL SUMMARY

**Data Scientist & Applied AI Engineer** with 3+ years of experience in machine learning, applied AI, and data-driven innovation, focused on building scalable systems for operational efficiency and decision-making. Skilled in designing and deploying end-to-end ML pipelines, advanced analytics, and real-time data solutions using Python, SQL, Spark, and AWS. Experienced in applying Generative AI and agentic workflows with LangGraph, RAG, and LLM orchestration, driving actionable insights and business optimization across financial services and real estate.

## EDUCATION

### EMORY UNIVERSITY

Master of Science in Business Analytics

Atlanta, GA

May 2025

### UNIVERSITY OF ZIMBABWE

Bachelor of Business Studies and Computing Science

Harare, Zimbabwe

Dec 2021

## PROFESSIONAL EXPERIENCE

### DATA SCIENTIST CONSULTANT | Pennybacker Capital - Austin, Texas

Dec 2024 - May 2025

- Designed and deployed machine learning models to forecast quarterly Gross Asset Value (GAV) for a \$4B+ real estate portfolio achieving 1% forecasting error (MAPE) and helping prevent an estimated \$2M in annual losses.
- Engineered an integrated data pipeline, consolidating over 50 internal and external data sources on Databricks to create a comprehensive datasets for predictive modeling and analysis.
- Initiated and analyzed Google Reviews data to create a sentiment-driven early warning system, identifying operational risks and opportunities for improvement across 11 underperforming multifamily properties.
- Translated complex model predictions into actionable business strategy by using SHAP and LIME to interpret feature importance and predictive insights enhancing stakeholder trust and data-driven decision-making.

### SENIOR DATA SPECIALIST | AFC Commercial Bank - Harare, Zimbabwe

Mar 2024 – Jun 2024

- Led the partnership between OK-Supermarket and AFC Bank for the OK Grand Challenge promotion, driving data-driven marketing strategies; effort generated a 200% increase in POS transactions across 70+ outlets.
- Developed a data visualization dashboard using Python, Apache Spark and Dash Plotly to analyze 20000+ ATM and POS terminal activity, providing critical insights and facilitating in-depth analysis and swift resolution of operational issues.
- Implemented an XGBoost model to predict point-of-sale client churn, enhancing targeted retention campaign effectiveness by 25% and reducing churn rates by 15% within two months.
- Led customer and loan data migration from T24 to IDC Core Banking System, achieving 99.4% accuracy by automating workflows with Python and Apache Spark for faster data cleaning and validation while minimizing downtime.

### DATA SPECIALIST | AFC Commercial Bank - Harare, Zimbabwe

Jun 2022 – Feb 2024

- Developed a Python backend with FastAPI to integrate the Reserve Bank of Zimbabwe (RBZ) API for the Credit Reference Bureau (CRB), reducing data processing time by 40% while enhancing regulatory compliance.
- Built ETL data pipelines using Apache Kafka and Python to integrate data from the core banking system, delivering accurate KPIs across 45 AFC Commercial Bank branches.
- Redesigned and optimized merchant reporting services with Apache Airflow and DBT, automating manual processes and increasing efficiency by 80%, while delivering insights on transaction performance to key stakeholders.
- Modernized a monolithic reconciliation app into scalable microservices using Docker, Python, FastAPI, Kubernetes and Angular, boosting efficiency by 150%.

## SELECTED PROJECTS

### QueryCraft AI Project - [github](#)

Jun 2025

- Developed QueryCraft-AI, a full-stack application that empowers users to construct complex data queries using a conversational AI powered by Python and LangGraph. The platform has a Next.js frontend, Fast API Backend and is built for production with Docker, LangFuse for LLM observability, MCP for integrating with external tools, and ZITADEL for secure user authentication.

### Sports Ticket Sales Forecasting - [github](#)

Feb 2025

- Achieved 3.3% forecast error in predicting Atlanta Braves ticket sales, the best in the competition. Developed XGBoost and LSTM models, integrating attendance, promotions and weather data to enhance forecasting precision.

### Log-Realtime-Analysis Project - [github](#)

Dec 2024

- Designed a scalable architecture for real-time log processing and visualization, handling 60,000 log events per second using a Kafka-Spark ETL pipeline, DynamoDB for real-time metric storage, and Python Dash for interactive dashboards.

TECHNICAL CAPABILITIES

---

Programming Languages	: Python, Java, R,PHP, SQL
Machine Learning & AI	: TensorFlow, PyTorch, scikit-learn, Darts, LLMs, RAG, LlamaIndex, LangGraph, MCP
Data Engineering	: Apache Spark, Kafka, Apache Airflow, Docker, Kubernetes, dbt, S3
Visualization & Reporting	: Dash (Plotly), Power BI, Tableau, Excel, Matplotlib, Seaborn
Cloud & Databases	: AWS, Databricks, PostgreSQL, SQL Server, DynamoDB, MongoDB, DuckDB, Redshift, NoSQL
Certifications	: Azure AI Fundamentals, Databricks: Generative AI Fundamentals, Oracle: Java 11 SE Developer