

# UNARINE TIMOTHY TSHIMAUWU

Data Scientist | ML Engineer | Data Analyst

Gauteng, South Africa | 082 580 4035 | [timothytshimauwu@gmail.com](mailto:timothytshimauwu@gmail.com)

LinkedIn: [utshimauwu](#) | Portfolio: [TimothyTshimauwu](#)

## PROFESSIONAL SUMMARY

Data Scientist and Analytics professional with 5 years of engineering and analytics experience, including 3+ years delivering predictive models, risk analytics, and BI solutions across financial services and large-scale manufacturing. Strong background in Python, SQL, Azure, and Power BI with documented contributions to R83M+ in cost savings, operational improvements, and revenue retention. Experienced in end-to-end analytics pipelines, stakeholder-driven reporting, and production-grade ML architecture. Seeking Data Scientist, ML Engineer, and Business Intelligence roles in banking, finance, and operations.

## CORE COMPETENCIES

- **Analytics & Statistics:** Predictive Modeling | Statistical Inference | Time-Series Analysis | Regression Analysis | Hypothesis Testing
- **Programming & BI:** Python | SQL | R | Power BI | DAX | Tableau | Excel | SAS (foundational)
- **Machine Learning:** Classification | Regression | Feature Engineering | Model Validation | XGBoost | scikit-learn | Clustering (K-Means, PCA)
- **Data & Cloud:** ETL Pipelines | SQL Server | Azure Data Factory | Databricks | AWS (EC2, S3) | Docker | Data Governance
- **Soft Skills:** Project Management | Business Acumen | Stakeholder Engagement | Problem Solving

## PROFESSIONAL EXPERIENCE

**Financial Data Analyst** | Ruach Consulting Solutions | Pretoria, Gauteng

*December 2024 - Present*

- Developed SQL-based data extraction and Power Query transformation workflows processing 50,000+ monthly NGO transactions, automating previously manual reporting processes through scheduled Power BI refreshes. Reduced monthly reporting time from approximately 40 to 20 hours
- Built 12 Power BI dashboards for cash flow monitoring, expenditure tracking, and funding performance across donor streams
- Performed variance analysis on R4.5M expenditure, identifying budget leakages, delivering recommendations projected to achieve R150K annual savings (5% cost reduction)

**Process Engineer (Operations Data Analyst)** | ArcelorMittal South Africa | Vereeniging, Gauteng

*March 2022 - November 2024*

- Built Azure data engineering pipeline (Data Factory, SQL Server, Power BI) processing 250,000+ monthly production records. Conducted bottleneck analysis identifying 6 critical equipment constraints. Cross-functional implementation contributed to R38M in operating profit through an 8% increase in throughput and a 5% reduction in downtime
- Conducted time-series trend analysis on 5 years of customer complaint data, identifying equipment failure patterns correlated with maintenance schedules. Findings contributed to corrective actions that retained R30M+ annual-value customer account
- Developed a regression-based optimization model analyzing energy consumption patterns for a 1.8M tonne/year plant. Analysis identified operational adjustments achieving R15M in annual energy cost savings through reduced gas consumption

**Graduate Engineer** | ArcelorMittal South Africa | Vereeniging, Gauteng

*March 2020 - February 2022*

- Analyzed 180,000+ production and quality records using SQL and Python, supporting cross-functional initiatives that delivered an 8% throughput improvement and a 10% reduction in customer complaints
- Optimized hot rolling mill process control using 50,000+ equipment records, contributing to R15M in mechanical-loss savings through improved operational efficiency and reduced unplanned downtime

## KEY ANALYTICS PROJECTS

### Credit Default Risk Management System | Portfolio Project | [GitHub Repository](#)

Python | SQL | SAS | XGBoost | FastAPI | Streamlit | Power BI | AWS

- **Business Problem:** Bank needed scalable credit scoring system to reduce default rates and optimize lending decisions
- **Approach:** Developed XGBoost classification model on 1M+ loan application dataset with 45 engineered features (DTI ratio, credit utilization, payment patterns), achieving 87% precision and 82% recall through hyperparameter tuning and 5-fold cross-validation
- **Outcome:** Built FastAPI scoring service containerized with Docker and deployed to AWS EC2. Architecture demonstrates production-ready design with error handling, logging, and health checks. Projected 15% bad debt reduction based on test set performance vs baseline approval strategy

### ArcelorMittal Production Analytics & Azure Pipeline | Arcelor Mittal experience | [GitHub Repository](#)

Python | Azure Data Factory | SQL Server | Power BI | Statistical Analysis

- **Business Problem:** 504,000 tonne/year steel plant missing production targets due to limited visibility into equipment-level bottlenecks
- **Approach:** Built an Azure data engineering pipeline using Data Factory to automate ETL, reconstructed equipment cycle times from event data, and applied Pareto analysis on downtime with Power BI dashboards
- **Outcome:** Identified six critical constraints. Methodology aligned with an ArcelorMittal initiative that delivered R38M in operating profit through an 8% throughput increase, 5% downtime reduction, and the first 40kt/month production milestone

### Call Center Performance Dashboard & Operational Analytics | Portfolio Project | [Tableau Public](#)

SQL | Tableau | Excel | Data Cleaning | Operational Analytics

- **Business Problem:** Call center operations lacked visibility into employee performance, call patterns, and service quality metrics needed for data-driven management decisions
- **Approach:** : Analyzed 5,000+ call center interactions tracking 8 operational KPIs (call volume, speed of answer, resolution rates, agent performance, satisfaction). Built interactive Tableau dashboard with time-based filters and agent drill-downs
- **Outcome:** Identified three underperforming agents, exposed a Monday staffing gap with a 26% call-volume spike, and highlighted speed-of-answer optimization opportunities. Insights supported targeted interventions projected to lift resolution rates to 75%, cut average handle time by 10%, and deliver R1.5M in annual savings through reduced escalations and improved staffing.

### UFood Marketing Campaign Optimization & Customer Segmentation | Portfolio Project | [GitHub Repository](#)

Python | Pandas | Scikit-learn | K-Means | Marketing Analytics | Statistical Analysis

- **Business problem:** Brazil's leading food delivery app saw flat profit growth due to inefficient mass marketing. Campaign acceptance averaged 15%, driving poor ROI.
- **Approach:** Analyzed 1,843 customers across 29 behavioral and transactional features. Applied K-Means clustering identifying 4 segments with 8.8x variance in campaign acceptance (8% to 73%). EDA revealed household size as key driver, customers with 0-1 children showed highest engagement and spend
- **Outcome:** Designed precision targeting strategy focused on top 50% of customers generating 91% of revenue. Projected 3.4x acceptance improvement (15% to 52%) and 71% cost reduction by eliminating catalog distribution to non-responsive segments

## EDUCATION

### Post Graduate Diploma in Data Science (Cum Laude) | University of the Witwatersrand | 2025

- Specialization: Machine Learning, Programming (Python, R), Mathematical Foundations

### BEng (Hons): Industrial and Systems Engineering | University of Pretoria | 2023

- Specialization: Enterprise Engineering, Optimization, Manufacturing Planning & Control

### BEng: Metallurgical Engineering | University of Pretoria | 2019

- Specialization: Process Control, Calculus, Statistical Analysis

*References available upon request*