

Ibrahima Barry

Data Scientist

Brooklyn, New York | iblouse49@gmail.com | 614-312-4440 | LinkedIn | GitHub | Portfolio

Summary

- Data Scientist with 5+ years delivering analytics and machine learning solutions from problem definition through deployment.
- Built forecasting, anomaly detection, and predictive models that improved conversion and retention and generated \$400K+ in quantified annual savings.
- Experienced translating business questions into measurable analyses and presenting results through dashboards and written recommendations.

Skills

- **Programming:** Python, SQL (Redshift, BigQuery), PySpark, R, PyTorch, TensorFlow
- **Machine Learning and AI:** XGBoost, Random Forest, Neural Networks, Ensemble Methods, K-means Clustering, Time Series Forecasting, LLM Integration, Retrieval Augmented Generation
- **Cloud and Tools:** AWS (S3, Glue, Redshift, SQS), Google Cloud (BigQuery, Vertex AI), Databricks, Snowflake, MLflow, Docker
- **Statistics:** Univariate and bivariate analysis, chi-square tests, ANOVA, summary statistics
- **Visualization:** Tableau, Power BI, SAS Visual Analytics, executive dashboards, KPI reporting
- **Languages:** French (fluent)

Experience

Senior Data Science Consultant, Data In Motion | New York, NY | Nov 2024 to Present

- Led an analytics engagement from problem framing through deployment; improved conversion rates by 18% using funnel and survey analysis.
- Built time series forecasting models in Python to optimize webinar scheduling; improved resource allocation efficiency by 20%.
- Delivered competitor analysis across 8 platforms; contributed to a 15% improvement in customer retention.

Analytics Engineer, USA for UNHCR | New York, NY | May 2022 to Jan 2024

- Implemented anomaly detection on donation approvals with automated alerting using Python and AWS SQS; prevented fraudulent transactions and saved \$400K annually.
- Built a production Redshift pipeline with AWS Glue integrating Salesforce and S3; wrote complex SQL across multi-million row datasets and reduced manual processing by 35%.
- Developed XGBoost models with engineered features (normalization, polynomial transformations, rolling statistics) to predict donor likelihood; delivered a 20% lift in campaign donations.
- Performed statistical analysis including chi-square tests to quantify city campaign performance; identified drivers behind a 25% uplift in digital conversions.

- Applied k-means clustering for segmentation and built Tableau dashboards to guide targeted retention campaigns.
- Delivered dashboard-based insights to senior leadership to support retention decisions across departments.

Data Science Fellow, Sharpest Minds | New York, NY | Mar 2021 to Mar 2022

- Developed predictive models to assess Type 2 diabetes risk using CDC BRFSS data.
- Processed and cleaned 400,000+ rows and 275 columns; prioritized sensitivity and achieved 83.77% sensitivity with logistic regression.
- Identified 25 key predictors of Type 2 diabetes including BMI, general health, and income level.

Data Analyst, Self-Employed | New York, NY | Feb 2013 to Feb 2021

- Forecasted seasonal product demand using ARIMA; improved winter sales by 15% and reduced unsold stock by 10%.
- Identified a 30% demand spike on Fridays and Saturdays; increased weekly avocado orders by 25% and moved deliveries to Thursday, resulting in a 12% sales boost and fewer stockouts.

Manager, Morning View Care Center | Columbus, OH | Jan 2003 to Jun 2006

- Trained and onboarded staff and provided coaching to support consistent service delivery.
- Coordinated and scheduled client medical appointments to improve continuity of care and timely access to services.
- Managed client financial transactions including bank deposits and bill payments, maintaining confidentiality and compliance.
- Built staff schedules to improve coverage and reduce gaps in service.
- Delegated daily responsibilities and monitored completion to strengthen accountability.

Projects

Proactive Churn Prevention System | Python, AI Agents, Lifelines, Vertex AI | Repo | Medium | Portfolio

- Built a dual-model architecture (classification plus survival analysis) to predict churn timing; identified \$2.54M in at-risk CLV and derived an intervention window (Day 45 to Day 95).
- Validated interventions through 5-variant A/B testing with Bonferroni correction; achieved 54.4% churn reduction ($p < 0.0001$) and 158.8x ROI across channels.
- Developed a multi-agent system using Google ADK to operationalize monitoring, risk scoring, and channel selection.

Multi-Agent Loan Processing System with RAG Underwriting | Python, Gemini, Transformers | Repo | Medium | Kaggle

- Designed a multi-agent loan processing workflow including KYC verification, financial data extraction, risk scoring, and RAG-based underwriting.
- Generated synthetic financial data for privacy-compliant demonstrations and implemented financial calculations for realistic loan assessments.
- Demonstrated end-to-end deployment of an LLM-based business application.

Fraud Detection Model and MLOps Deployment | Python, MLflow, Streamlit | Repo

- Built and deployed an end-to-end model for fraudulent transaction detection from EDA through production deployment.
- Integrated MLflow for experiment tracking and Streamlit for a real-time fraud detection interface.
- Built a user-facing application demonstrating full model lifecycle management.

Mock-Interview App | Python, Generative AI, Streamlit | App

- Built an interactive Streamlit chatbot that simulates job interviews with dynamic question generation based on role and company context.
- Integrated an LLM for conversational interview flow and automated feedback scoring.
- Deployed a user-facing GenAI application with a production-style interface.

Education

Master of Statistics and Applied Mathematics, CUNY Hunter College | New York, NY

Certifications

- AWS Certified Machine Learning - Specialty (Mar 2024)
- AWS Certified Cloud Practitioner (Feb 2024)
- SAS Certified Clinical Trials Programmer Using SAS 9 (Dec 2020)
- SAS Certified Specialist: Forecasting and Optimization Using SAS Viya 3.5 (Sep 2020)
- SAS Certified Specialist: Machine Learning Using SAS Viya 3.4 (Aug 2020)
- SAS Certified Base Programmer for SAS 9 (Aug 2020)