

Abba Otieno Ndomo

Atlanta, GA | +1 (404) 819-5945 | abba.ndomo@gmail.com | Portfolio

Data Analyst / Data Scientist Intern

Work Authorization: F-1 (CPT eligible) | OPT (May 2027) | STEM OPT eligible | No sponsorship required for internship (CPT)

Technical Skills

Languages: SQL (CTEs, Window Functions, Joins), Python (Pandas, Scikit-Learn), R, Bash

Tools & Ecosystem: Snowflake, Airflow, Git (Version Control), Tableau, Excel (VBA/Macros), dbt, VS Code

Analytics: A/B Testing, Experiment Design, KPI Definition, ETL Pipelines, Hypothesis Testing, Regression Analysis

Experience

KPMG (Automation/NLP)

Mumbai, India

Software Engineering Intern (Data & Analytics Focus)

Jun 2024 – Aug 2024

- Manual audit QA processes were prone to human error and high latency.
- Engineered a RAG-based NLP pipeline using Python and SQL to automate document verification; optimized data retrieval with advanced SQL indexing.
- Achieved 95% precision in automated audits and improved data pipeline test coverage by 40% using a scalable PyTest suite.

GT Investment Committee

Atlanta, GA

Quantitative Analyst Mentee

Jan 2025 – Present

- Manual evaluation of ticker performance delayed rebalancing of a \$500K ETF portfolio.
- Developed an automated equity screener in Python using `pandas` and `yfinance`; built regression models to identify alpha in undervalued assets.
- Reduced manual research time by 60% and provided data-driven buy/hold recommendations adopted by the committee chair.

Selected Projects

Boxxer AI — ML-Powered Packing Assistant

Jan 2024 – Present

React Native, Python (Flask), Knapsack Optimization, FastAPI, SHAP

- Designed and deployed travel assistant app to optimize item recommendations within weight limits.
- Trained ML model on 300+ user sessions; implemented SHAP analysis to rank user-specific packing priorities.

Loan Default Predictor (End-to-End Analytics Case Study)

Independent Project

SQL, Python, XGBoost, SHAP

Nov 2024 – Dec 2024

- Needed to identify high-risk borrowers to minimize financial loss for a synthetic lending dataset.
- Executed complex ETL using SQL (CTEs, Null handling) to clean raw data; trained an XGBoost classifier and applied SHAP for model interpretability.
- Delivered a model with 94.4% AUC and a summary memo for "stakeholders" explaining the top 3 risk drivers (Debt-to-Income, Credit Score, History).

Queueing Analysis of Hemphill Intersection — QPLEX, Python, Data Analysis

Dec 2024

ISYE 3232, Georgia Tech

- Parsed 600+ vehicle events from intersection video footage to calculate interarrival, waiting, and service times.
- Modelled as M/M/1 queue; used QPLEX to simulate pedestrian disruption impact, showing 17.4% increase in average wait.

Quantitative Trading Library

Jan 2024 – May 2024

Python, Black-Scholes, Monte Carlo

Atlanta, GA

- Needed a scalable framework to price complex derivatives and backtest strategies.
- Built an object-oriented library in Python implementing Black-Scholes and Monte Carlo simulations; integrated LOB dynamics for backtesting.
- Developed high-performance delta-hedging logic that improved risk-adjusted return simulations for options spreads.

Education

Georgia Institute of Technology

Atlanta, GA

BS, Industrial & Systems Engineering; Minor in Mathematics

Aug 2023 – May 2027

- Relevant Coursework: Probability & Statistics, Linear Algebra, Optimization, Data Engineering Operations.

Awards and Honors

Dean's List @ Georgia Institute of Technology

Peer Inspiration Award @ Georgia Tech ISyE Center for Academic Success

Third Place – OKSEF International Engineering Fair (Chemical Eng. Category)

Leadership & Engagement

Chair – DEI Committee, ISyE Department

Atlanta, GA

Analytics Working Group Lead

2023 – 2024

- Analyzed climate survey data from 500+ respondents; synthesized key insights into a visual dashboard presented to department heads.
- Drafted an action plan addressing top student concerns; recommendations were officially adopted into the FY24 program strategy.