Arnold Muzarurwi

Jersey City, NJ | amuzarur@mail.yu.edu | (201)-423-0839 | LinkedIn | GitHub

SUMMARY

Data analyst skilled in SQL, Python (pandas), Tableau, and Excel, experienced in cleaning and linking messy data, building dashboards, and writing short findings for decision makers. Comfortable turning ambiguous business questions into concrete analysis for clients and non-technical audiences.

EDUCATION

Yeshiva University, Katz School of Science and Health — New York, NY
Master of Science in Data Analytics and Visualization, GPA 3.7 — Expected December 2026

Beijing Institute of Technology — Beijing, China
Bachelor of Science in Computer Science and Technology, GPA 3.6 — June 2024

PROJECTS

NYC Rideshare Driver Pay and Congestion Pricing Analysis — October 2026
MySQL, SQL, Python, Pandas, Statsmodels, Tableau, Git

- Built a MySQL star schema (trip fact table with company and zone dimensions) from NYC TLC Uber and Lyft trip records, loading a 3.55M trip sample from Feb to Apr 2023, 2024 and 2025 and flagging 1,547 invalid records through SQL data quality checks.
- Wrote SQL KPI queries on driver share of fare, trip speed, rider fares and driver pay per hour, comparing trips that start or end in the 37 congestion pricing zones against all other trips before and after the policy.
- Measured the policy with difference in differences regressions and a 2023 placebo test: zone trips sped up 0.79 mph more than other trips on Uber (0.41 on Lyft) with no earlier trend, while a $2.25 rise in Uber zone fares proved to be a trend that predated the fee.
- Built and published an interactive Tableau Public dashboard with KPI cards, a zone map of driver pay per hour, an hour by weekday pay heatmap and a fee year versus prior year comparison, presenting the findings for a general audience.

Financial Fraud Detection — May 2026
Python, Neo4j, Cypher, GDS, Pandas

- Modeled 6.3M mobile money transactions across 9M+ accounts as a 15.4M-node graph to detect fraud patterns invisible at the individual-transaction level.
- Showed the existing rule-based system caught only 16 of 8,213 fraud transactions, a 99.81% miss rate concealing over $12B in fraudulent transfers.
- Identified 2,660 fraud account pairs unsupervised via Louvain community detection (about 65% recall) and confirmed sink accounts through convergence of three independent methods.

NYC Street Tree Equity Analysis — May 2025
Tableau, Excel

- Analyzed the 2015 NYC Street Tree Census with 2024 neighborhood income data in Tableau for a nonprofit sponsor (New York Restoration Project) to surface environmental equity gaps in urban tree coverage.
- Built Tableau visualizations (borough comparisons, a bubble chart, and a choropleth income map) showing Queens holds 36.89% of street trees versus Manhattan's 9.36%, and that tree coverage does not track neighborhood income.
- Translated findings into targeted, equitable tree-planting recommendations for a non-technical stakeholder audience.

PROFESSIONAL EXPERIENCE

IT Support Intern — August 2024 to December 2024
Avana Motors, Harare, Zimbabwe

- Tracked and analyzed incident records in a ticketing system to monitor service trends, support SLA compliance, and inform data-driven improvements to service quality.
- Applied structured root-cause analysis to diagnose recurring system and connectivity issues, documenting findings and escalating complex cases to senior technicians.
- Audited Microsoft 365 configurations and endpoint settings across the organization to identify access issues and maintain reliable day-to-day operations.

SKILLS

Programming & Query: SQL, Python, R, Cypher
Analysis & Reporting: Data Cleaning, Data Wrangling, Exploratory Data Analysis, Statistical Modeling, Regression Analysis, Causal Inference, KPI Reporting, Feature Engineering
Visualization & Reporting: Tableau, Excel, Matplotlib
Databases: MySQL, Neo4j
Tools: Git, ETL, Data Warehousing
Certifications: Neo4j Graph Data Science Certification (Neo4j GraphAcademy, Apr 2026)