Samuel Okotako

Data Engineer

Catania, Italy | +39 351 092 0026 | francissamuel840@gmail.com
Portfolio · GitHub · LinkedIn

PROFESSIONAL SUMMARY

Higher Technical Diploma holder in Data Management building Python and SQL pipelines for ingestion, incremental transformation and quality checks. My projects use DuckDB and dbt for validated analytical models, and PySpark on Databricks for raw-data ingestion. I apply testing and traceability to relational data.

PROFESSIONAL EXPERIENCE

Spindox | Data Management Intern Catania, Italy | June–October 2026

- Studied incremental ingestion, relational SQL and Databricks during internship technical training.
- Developed an individual *Data Privacy Governance* project on synthetic data, inventorying 109 PostgreSQL columns and implementing masking and pseudonymization transforms.
- Validated integrity and relationships across 20 protected tables; documented 182 passing tests across the full suite.

SELECTED PROJECTS

Incremental e-commerce data pipeline Personal project | Python, DuckDB, dbt | Case study

- Built a Python, DuckDB and dbt pipeline on synthetic orders; deterministic upserts handle 10 late inserts and 5 corrections without duplicate keys.
- Validated 5,010 current rows, idempotent replays and full-refresh parity; exposed sales marts and KPIs in a read-only Streamlit dashboard.

Inside Airbnb on Databricks ITS group project | PySpark, Delta

- Coordinated the group and wrote Bronze ingestion for four Amsterdam and Rio CSV files, adding city and ingestion timestamp metadata.
- Stored raw data as Delta tables; the group project also includes Silver transformations and Gold SQL views for comparative analysis.

TECHNICAL SKILLS

Data Engineering Python, SQL, ETL/ELT, dbt, DuckDB, Databricks

Data Platforms PostgreSQL, PySpark, Delta Lake, Bronze/Silver/Gold architecture

Cloud & DevOps AWS (ITS coursework), Docker, Git

Data Quality automated tests, validation, integrity, incremental pipelines

EDUCATION

ITS Steve Jobs Academy 2024–2026
Higher Technical Diploma in Data Management – EQF Level 5 | Final grade: 110/110

University of Catania
BSc in Computer Science (L-31) – ongoing

I.T. Archimede 2018–2023
Technical High School Diploma in Electronics and Electrical Engineering – Automation specialization

I authorize the processing of personal data in this CV for recruitment purposes in accordance with Regulation (EU) 2016/679.