Samuel Okotako

Data Engineer

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

PROFILO PROFESSIONALE

Tecnico Superiore Data Manager: sviluppo pipeline Python e SQL per ingestion, trasformazioni incrementali e controlli di qualità. Nei progetti uso DuckDB e dbt per modelli analitici verificabili e PySpark su Databricks per acquisire dati grezzi. Applico test e tracciabilità ai dati relazionali.

ESPERIENZA PROFESSIONALE

Spindox | Data Management Intern Catania | giugno–ottobre 2026

- Approfondito caricamenti incrementali, SQL relazionale e Databricks durante la formazione tecnica del tirocinio.
- Sviluppato il progetto individuale *Data Privacy Governance* su dati sintetici: inventariate 109 colonne PostgreSQL e implementate trasformazioni di masking e pseudonimizzazione.
- Verificate integrità e relazioni in 20 tabelle protette; documentati 182 test superati nella suite complessiva.

PROGETTI SELEZIONATI

Pipeline incrementale per ordini e-commerce Progetto personale | Python, DuckDB, dbt | Case study

- Costruita su dati sintetici una pipeline Python, DuckDB e dbt: l'upsert deterministico gestisce 10 inserimenti tardivi e 5 correzioni senza duplicati.
- Verificate 5.010 righe correnti, replay idempotenti e parità con il full refresh; mart e KPI sono consultabili in una dashboard Streamlit.

Inside Airbnb su Databricks Progetto ITS di gruppo | PySpark, Delta

- Coordinato il lavoro di gruppo e scritto il notebook Bronze per acquisire quattro CSV di Amsterdam e Rio, aggiungendo città e timestamp di ingestion.
- Salvati i dati grezzi come tabelle Delta; il progetto di gruppo comprende trasformazioni Silver e viste SQL Gold per analisi comparative.

COMPETENZE TECNICHE

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

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

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

Data Quality test automatizzati, validazione, integrità, pipeline incrementali

FORMAZIONE

ITS Steve Jobs Academy 2024–2026
Diploma di Tecnico Superiore Data Manager (EQF 5) – 110/110

Università degli Studi di Catania
Laurea in Informatica (L-31) – In corso

I.T. Archimede 2018–2023
Diploma in Elettronica ed Elettrotecnica – articolazione Automazione

Autorizzo il trattamento dei dati personali contenuti nel presente curriculum vitae per finalità di ricerca e selezione del personale, ai sensi del Regolamento (UE) 2016/679.