MAHMOUD HESHMAT

DATA ENGINEER

Cairo, Egypt | mheshmat566@gmail.com | +20 111 826 6047
linkedin.com/in/mahmoud-heshmat-2455022b5/ | github.com/Mahmoud336u

PROFESSIONAL SUMMARY

Data engineer with freelance experience building Spark ETL pipelines, Airflow workflows, and PostgreSQL data models for e-commerce and telecom analytics. Develops Python API ingestion services and deploys applications on AWS. Focused on reliable batch processing, SQL transformations, data quality, and maintainable delivery.

TECHNICAL SKILLS

Languages and processing: Python, SQL, Apache Spark/PySpark, ETL/ELT, REST API ingestion

Orchestration and quality: Apache Airflow, Great Expectations, retries, backfills, validation, monitoring

Databases and modeling: PostgreSQL, MongoDB, dimensional modeling, star schemas, incremental UPSERTs, SQL window functions

Cloud and delivery: AWS S3, RDS, ECS Fargate, ECR, Kafka/MSK, Docker, GitHub Actions, Terraform, pytest

RELEVANT CLIENT EXPERIENCE

Data Analyst | Freelance | Nov 2024 - Feb 2025

E-commerce and telecom clients | Data pipelines and analytics

- Built three Spark ETL pipelines and repaired a fourth, integrating four CRM/ERP sources and processing approximately 100 GB per daily run using AWS S3, Kafka/MSK, and Airflow.
- Designed a PostgreSQL star schema, incremental UPSERTs, and SQL window-function transformations to prepare data for downstream analytics.
- Reduced a core 100 GB batch runtime from 4.5 to 1.5 hours on comparable workloads, shortening daily data-processing time by approximately 67%.
- Orchestrated daily Airflow runs with retries, failure alerts, and backfills; applied Great Expectations validation to support dependable pipeline operation.
- Delivered Tableau dashboards and Python/SQL customer segmentation and churn analysis; handed over DAGs, runbooks, and validation logs to client IT/Data Ops teams.

ADDITIONAL ENGINEERING EXPERIENCE

AI/ML Engineer | Independent Product and AI Engineering | Mar 2025 - Present

Meridian Personal Finance Assistant | Personal/beta MVP

- Built Python/FastAPI services and Plaid Sandbox ingestion for an application supporting test-account linking, transaction categorization, and monthly spending forecasts.
- Prepared approximately 75,000 sandbox and synthetic transaction records for model development, using transaction descriptions and merchant category codes with chronological train/test separation.
- Deployed the application on AWS ECS Fargate and packaged a reproducible local environment with Docker Compose.
- Automated pytest checks, Docker builds, and ECR publishing with GitHub Actions; managed infrastructure with Terraform and monitored services with Prometheus.

SELECTED PROJECT

AI Multi-Agent Platform | Independent project

- Implemented PostgreSQL querying and vector retrieval with RDS PostgreSQL/pgvector, HNSW indexing, and metadata filtering.
- Added schema mapping, few-shot SQL examples, and bounded query-correction retries to a LangGraph agent; evaluated retrieval quality on 50 curated query-context pairs.

CERTIFICATIONS

AWS Certified Machine Learning Engineer - Associate | 2026

Google Data Analytics Professional Certificate | MLOps Specialization, Duke University

EDUCATION AND LANGUAGES

Bachelor of Engineering, Computer Engineering | El Shorouk Academy | 2026

Arabic: Native | English: Professional working proficiency | German: B1