Elias Velazquez

Data Engineer AWS

Buenos Aires, Argentina
Email | Website | LinkedIn | GitHub

PROFESSIONAL SUMMARY

Data Engineer specializing in cloud-native data platforms, high-throughput pipelines, and event-driven architectures on AWS and Snowflake. Proven track record of scaling production systems to handle **1.5 TB+ of monthly ingestion** and **25x YoY daily volume growth**, while maintaining resilient, cost-effective architectures. Expert at optimizing heavy compute workloads—slashing warehouse runtimes by **85%** via `write_pandas`—and decoupling high-traffic bottlenecks using asynchronous messaging (SQS). Experienced in Python-driven ETL/ELT, multi-tenant isolation, and programmatic data warehouse governance.

CORE SKILLS

- Languages: Python (Advanced: OOP, Asyncio/Threading, Type Hinting, Structured Logging), SQL (CTEs, Window Functions, Joins), Bash.
- AWS Ecosystem: Lambda, S3, ECS/Fargate, SQS, EventBridge, DynamoDB, Step Functions,API Gateway, IAM, CloudWatch.
- Data Engineering & Tools: Snowflake, Pandas, Boto3, Docker, Git, ETL/ELT Pipelines,Multi-Tenant Architecture, Schema Evolution.
- DataOps & DevOps Foundations: GitHub Actions CI/CD, Terraform (IaC foundations), Linux Automation.

PROFESSIONAL EXPERIENCE

Data Engineer — Vordentech

February 2024 – Present

- **Pipeline 1: Multi-Tenant Legacy Data Processing Engine**
  - Designed and deployed a containerized, event-driven processing platform (ECS/Fargate, S3, EventBridge) to ingest and parse 4-5GB fixed-width legacy financial files containing up to 5M rows per execution.
  - Engineered a Python/Pandas streaming and chunk-based pipeline (100k rows/chunk) to isolate and generate multi-tenant datasets for up to 13 separate tenants concurrently without memory overhead.
  - Developed custom position-aware parsers to handle complex COBOL-style specifications, non-standard date encodings, and signed numerics, reducing legacy service runtime by ~35% through code refactoring and structural optimization.
- **Pipeline 2: High-Performance Ingestion & Storage Architecture**
  - **Architected a schema-driven AWS Lambda engine** that infers schema evolution and auto-provisions Snowflake tables, scaling seamlessly to support **25x YoY daily volume growth** (from ~2 GB to 51+ GB/day).

- **Engineered pipeline resilience** to absorb a **65% MoM surge**, successfully eclipsing **1.5 TB of monthly ingestion** while maintaining strict stability and cost-efficiency.
- **Slashing ETL runtimes by ~85%** (from 15–20 mins down to 2–3 mins) by replacing native Python writes with the Snowflake Connector's `write_pandas` method.
- **Designed an asynchronous metrics pipeline** where Pipeline 2 offloads execution metadata to AWS SQS, decoupling processing from storage to eliminate Snowflake warehouse write-saturation and ensure ordered inserts under heavy traffic.
    - **Implemented DataOps best practices** (Docker, structured logging, alerts) and collaborated with global stakeholders to align architecture with production needs.

EDUCATION

Technical Degree in Data Science & AI — Universidad Nacional de Chilecito (March 2026 – Present)

- Coursework: Data Analysis, Machine Learning Fundamentals, Statistics, and Mathematical Foundations.

CERTIFICATIONS

- AWS Certified Solutions Architect Associate (Issued Sep 2026 - Expires Sep 2029)

LANGUAGES

- English: C1 (Advanced/Fluent - Professional Working Proficiency)
- Spanish: Native