David Damon

Tampa, FL • (316) 609-8566 • dddamon06@gmail.com
[LinkedIn](#) | [GitHub](#) | [Portfolio](#)

Data Engineer

ETL/ELT · Cloud Platforms (Azure, GCP) · Databricks/Spark/Delta · dbt · Monitoring · CI/CD

Profile

Data Engineer delivering production-grade lakehouses and analytics platforms across cloud environments. Experienced in orchestrating ETL/ELT pipelines, building medallion architectures, applying data quality frameworks, and operationalizing deployments with CI/CD and monitoring. Strong focus on clean architecture, profiling, config-driven pipelines, and clear documentation to ensure scalability and reliability.

Core Skills

- **Cloud & Orchestration:** Azure Data Factory, Airflow, Azure Monitor, ARM Templates
- **Storage & Compute:** ADLS Gen2, GCS, Databricks (Spark, Delta Lake), BigQuery, SQL Warehouse
- **Transform & Modeling:** PySpark, dbt (Core), SQL (window functions), Medallion architecture
- **Quality & Governance:** dbt tests (not_null/unique/accepted_values), Great Expectations, IAM principles
- **Ops & CI/CD:** GitHub Actions, structured logging, runbooks, cost/perf optimization, Docker
- **Programming & Interop:** Python 3.11 (OOP, Pandas, requests, typing), REST/JSON APIs

Selected Projects

Real-Time Offer Notification System — *Kafka + Spark Streaming + AWS (Production Ready)*

- Built production streaming pipeline processing **10K+ transactions/minute** with **P95 < 60 seconds** end-to-end latency.
- Implemented dual-sink pattern: S3 data lake (PII-hashed) + SQS notifications with **exactly-once processing**.
- Deployed on AWS with Lambda SMS delivery, Athena analytics, and **99.9% uptime** with fault tolerance.
- Designed secure architecture with **PII protection**, SASL_SSL encryption, and comprehensive monitoring.

E-commerce Data Warehouse — *GCP + BigQuery + dbt + GitHub Actions (Production Ready)*

- Built cloud-native medallion architecture processing **116K+ records** with **58% intelligent deduplication**.
- Implemented NULL-safe transformations, incremental MERGE models, and **18/19 comprehensive tests passing**.
- Deployed CI/CD pipeline with Workload Identity Federation; **complete audit trail** with 17,996 QA records.
- Generated interactive dbt documentation and **zero data loss** enterprise-grade reliability.

Earthquakes Lakehouse — *ADF → ADLS Gen2 → Databricks (Delta) + dbt, Azure Monitor*

- Parameterized ingestion (start/end date, min magnitude) from USGS into ADLS.
- Databricks Bronze → Silver → Gold refinement; dbt marts for daily metrics, top 100 quakes, dashboards.
- Enforced dbt tests in CI; daily scheduled loads; Monitor alerts + triage runbook.

Taxi Lakehouse — *ADF → ADLS Gen2 → Databricks (Delta) + dbt*

- Automated CSV ingestion to ADLS; Bronze → Silver transformation in Databricks; dbt marts for daily metrics.
- Daily 06:00 trigger, email alerts, validation toggles; optional ARM deployment.

Recruit Reveal — Low-Latency Model Serving on Databricks

- Feature pipelines (z-scores, encodings, missing-value flags).
- XGBoost models tracked in MLflow; deployed as SQL UDFs returning JSON predictions.
- Hardened with dtype guards; surfaced through a Next.js evaluation UI.

SEC EDGAR Financials Warehouse — *GCP + BigQuery + dbt + GE (Looker Studio)*

- Ingested SEC filings into a medallion architecture (sec_raw → sec_curated → sec_viz).
- Partitioned/clustered BigQuery marts cut scanned bytes by **80–90%**.
- 14 dbt tests, 100% GE validations; daily scheduled pipeline with smoke checks.

Cloud-Native Crypto ETL — *GCP + dbt + Terraform + GE (Looker Studio)*

- Cloud Run Jobs + Scheduler → GCS → BigQuery partitioned tables.
- Rolling analytics (7/30/90-day MAs, volatility, golden-cross).
- dbt + GE validations; **~$2.36/mo** query spend with bytes-scanned proof.
- Makefile-driven workflows with AI-assist tests.

Experience

React Engineer (Contract) — DevSoft, Tampa, FL | *Jun 2022 – Present*

- Built React apps with code-splitting, lazy-loading, and REST API integrations.
- Reduced median load time by **~40%** through performance optimizations.

Biology Teacher — Freedom High School (HCPS), Tampa, FL | *Aug 2023 – Present*

- Analyzed data for **150+ students**, producing visualization/statistical reports.
- Informed instruction for an 8-member team using data insights.

Education

Flatiron School — Full-Stack Web Development (Python/Flask & JavaScript), 2025 — Tampa, FL
Friends University — B.S., Biology, 2021 — Wichita, KS