HANUMANTHA RAO VULLI

Data Engineer | Python | SQL | AWS | Spark | Backend Systems

Denton, TX | vullihanumantharao@gmail.com
Portfolio: https://sunnyvulli.github.io/ | GitHub: https://github.com/Sunnyvulli
LinkedIn: https://www.linkedin.com/in/hanumantharaovulli/

PROFESSIONAL SUMMARY

Data engineer with AWS pipeline, SQL, and Java backend experience; MS in Data Engineering. Portfolio: Spark/Iceberg, dbt, Airflow, approval-gated LangGraph, MCP and reproducible retrieval evaluations.

TECHNICAL SKILLS

Languages and Backend: Python, SQL, Java; PostgreSQL, Oracle, SQLite; Spring Boot, REST APIs
Data and Agent Tooling: Spark, Iceberg, dbt, Airflow, LangGraph, MCP, Python unittest
Cloud and Delivery: AWS EC2/S3/RDS/Lambda/API Gateway; Terraform, Docker, Jenkins, GitHub Actions

PROFESSIONAL EXPERIENCE

Data Engineer | **PRC Cloud Tech, USA** | January 2026 - Present
- Support AWS ingestion, ETL, reporting SQL, validation, and Lambda/API Gateway/RDS services.

Data Engineer | **Tech Force, USA** | September 2025 - January 2026
- Supported AWS pipelines, data-quality checks, PostgreSQL/Oracle tuning, and Jenkins/Docker delivery.

Data Engineer | **Infosys Now, USA** | April 2025 - September 2025
- Worked on Spring Boot/AWS ingestion services, record validation, and workflow documentation.

Java Full Stack Developer | **DBS Technologies, India** | July 2022 - May 2023
- Developed Angular features and Spring Boot APIs; assisted AWS and Jenkins/Docker delivery.

SELECTED PROJECTS

Lakehouse Guardian | Spark, Iceberg, dbt, Airflow, LangGraph, Terraform
https://github.com/Sunnyvulli/lakehouse-guardian
- Processes 17,379 UCI Bike Sharing observations into local Iceberg snapshots; dbt dimensional models gate publication and approval is bound to a proposal hash.
- In 21 controlled fault batches, row detection precision was 1.00 and recall 0.80; documented missed plausible count inflation. Includes 13 Python tests and real Airflow execution in CI.
- Terraform S3/Glue configuration passes provider validation in CI; AWS infrastructure has not been deployed.

Catalog Sentinel | MCP, Python, SQLite, GitHub Actions
https://github.com/Sunnyvulli/catalog-sentinel
- Built two read-only MCP tools with host-assigned roles, strict schemas, bounded queries, and hashed audit records; 14 behavioral tests.
- Actual stdio evaluation enforced all 13 expected denials across 16 requests and left database bytes unchanged; local host trust, not remote authentication.

Evidence Ledger RAG | SQLite FTS5, scikit-learn, Python
https://github.com/Sunnyvulli/evidence-ledger-rag
- Incremental indexing and cited extracts; 18 tests. On 14 authored queries, hybrid expected-file hit@3 was 11/14 versus BM25 10/14. Documents missed files and semantic guard failures.

Mobility Quality Lab | Python, SQLite, Docker
https://github.com/Sunnyvulli/mobility-quality
- Includes 13 passing tests; 84-day UCI Bike Sharing backtest measured MAE 49.77 versus 64.03 rentals/hour for an online weekly-lag baseline.

Applied ML Research Suite | Reproducible baselines; 58 integrity tests
https://github.com/Sunnyvulli/Data-Engineer_AI
- Citation-support triage, phishing robustness, and simulated-transient ranking; measured outputs and failure analysis. Writing is unpublished.

EDUCATION

Master of Science in Data Engineering | University of North Texas | December 2024
Bachelor of Technology in Computer Science and Engineering | 2019 - 2022
KKR and KSR Institute of Technology and Sciences, Guntur

CERTIFICATIONS

AWS Certified Data Engineer - Associate | Oracle Certified Professional: Java SE 17 Developer