Aaron D. Gonzalez

Lead Data Engineer
Brooklyn, NY | 979.215.9052 | aaron@aarondgonzalez.com
aarondgonzalez.com | linkedin.com/in/aarondgonzalez | github.com/stubs

PROFESSIONAL SUMMARY

Lead Data Engineer with a decade of experience building and operating production data platforms on GCP and AWS. Owns migrations end to end, from a 500-model dbt Cloud to dbt Core move to a DataOps CI/CD rollout that cut Jenkins build times by 87%. Builds the self-service tooling, orchestration, and guardrails other teams depend on, and tunes BigQuery for cost.

EXPERIENCE

People, Inc.
New York, NY

Lead Data Engineer — Feb 2025 - Present

Senior Data Engineer — Oct 2023 - Feb 2025

- Project managed the DataOps rollout of CI/CD pipelines across the data organization, reducing Jenkins build times by 87%.
- Owned the organization's dbt Cloud to dbt Core migration end to end, moving approximately 500 models and tests and training the team on the dbt CLI workflow.
- Identified inefficiencies in BigQuery queries and implemented optimizations, reducing data scanned by 81% and contributing to a 22% reduction in reservation costs.
- Took an ambiguous, multi-repository Curator-to-Airflow migration from blank page to working system, writing the plan and sequencing the work across four codebases.
- Extended the platform so other teams could configure it without modifying core infrastructure, propagating run-time configuration into containers and downstream consumers and isolating dbt projects per environment.
- Shipped operational tooling that let other teams run their own recoveries: a tag-based DAG trigger interface, a historical backfiller, and a catchup DAG that takes user input for which pipelines to replay.
- Turned recurring production incidents into CI/CD build failures: stale branches, hardcoded full refreshes, and malformed YAML are now caught before merge.
- Cut mean time to detect pipeline failures by building dbt DAG alerting end to end and routing each alert to the team that owns it.
- Supervised direct reports and mentored junior engineers, running an ongoing Lunch & Learn series for technical and non-technical audiences.
- Adopted LLM-assisted development tooling and turned it toward defensive coding, using it to raise input validation and error handling standards across the team's repositories.

Density
New York, NY

Senior Data Engineer — Jul 2022 - Oct 2023

- Designed and implemented the data pipelines feeding the company's data lake and warehouses.
- Securely containerized data pipelines and deployed them on Amazon ECS and GCP Cloud Run.
- Introduced data observability the team previously lacked, configuring AWS CloudWatch metric alarms and SNS alerting.
- Managed the team's GCP project, designing and deploying the Data Analytics and Engineering group's cloud infrastructure.
- Configured custom IAM roles around the principle of least privilege to keep the team's cloud footprint secure.
- Applied the data modeling strategy that fit each use case, working in both Kimball dimensional and Inmon normalized styles.

Dotdash Meredith New York, NY
Senior Data Engineer Dec 2021 - Jul 2022
Data Engineer Dec 2017 - Dec 2021

- Interfaced directly with Data Science, Performance Marketing, and other teams to build custom pipeline solutions against their requirements.
- Maintained, optimized, and developed new data integrations curating the Presto and BigQuery data lake.
- Built the Optimistic Partition Adder, reducing a recurring task's execution time from over 20 hours to seconds.
- Refactored a legacy PySpark job, achieving a 92% reduction in execution time.
- Conducted design and code reviews with an emphasis on teaching junior team members.
- Reduced the introduction of technical debt in key repositories by putting previously unused Git hooks to work.

SKILLS

Languages: Python, SQL, Bash, Rust
Transformation & Modeling: dbt, Dimensional Modeling, Normalized Modeling, Pandas, PySpark, Apache Beam
Orchestration & Pipelines: Apache Airflow, Prefect, Fivetran, ELT, ETL
Data Platforms & Warehouses: BigQuery, Hive, DuckDB, Snowflake, Presto
Cloud & Infrastructure: Google Cloud Platform, Amazon Web Services, Terraform, Docker, Kubernetes
DevOps & Practices: CI/CD, GitHub Actions, Jenkins, Git, Linux, Data Observability, Looker

CERTIFICATIONS

Google Cloud Professional Data Engineer | dbt Analytics Engineering Certification | AWS Certified Developer

EDUCATION

Cornell University Ithaca, NY
B.A., American Studies
Brooklyn College, The City University of New York New York, NY
M.S., Education
Hunter College, The City University of New York New York, NY
M.S.W.