Benjamin Osei

Staff Data Engineer

benjamin.osei@email.com • (646) 555-0129 • New York, NY
linkedin.com/in/benjaminosei

Summary

Staff data engineer with 10 years in insurance, health analytics and e-commerce, setting data architecture for a marketplace with 9M shoppers. Designs streaming and batch platforms on Kafka, Spark and Snowflake, standardized dbt and Airflow across 5 teams and manages infrastructure in Terraform. Known for performance tuning that turns expensive jobs into cheap ones.

Experience

03/2022 - Present
New York, NY

Staff Data Engineer

Bowery Commerce

Staff engineer across 5 data teams at an online marketplace with 9M shoppers.

- Set the data architecture for **5** teams, publishing standards for dbt, Airflow, testing and ownership.
- Lead a Kafka and Spark streaming redesign that delivers order events to Snowflake in under 60 seconds.
- Cut Snowflake spend by $900K a year through performance tuning, clustering and warehouse scheduling.
- Manage Terraform modules for Snowflake, Fivetran and AWS used by 30 engineers.

05/2019 - 02/2022
New York, NY

Senior Data Engineer

Gotham Health Analytics

Senior engineer at a health data startup that grew from 40 to 250 staff.

- Built the first Airflow and dbt platform, growing it to 400 models and 120 DAGs.
- Designed HIPAA-compliant data models for 20M patient records across 60 hospital clients.
- Replaced 30 custom connectors with Fivetran, freeing 2 engineers for product work.
- Interviewed and onboarded 9 data engineers as the team scaled.

07/2016 - 04/2019
Melville, NY

Data Engineer

Long Island Mutual Insurance

Data engineer on the warehouse team of a regional insurer.

- Wrote Python and SQL ETL for policy, billing and claims data across 4 source systems.
- Moved nightly batch processing to Spark, cutting runtime from 9 hours to 3.
- Built data quality checks for 200 tables that flagged issues before morning reports.
- Tuned 40 slow SQL queries feeding actuarial reports, cutting their average runtime by 60%.

Skills

- Python
- SQL
- Airflow
- dbt
- Snowflake
- Spark
- Kafka
- Terraform
- Fivetran
- Performance Tuning

Education

05/2016
New Brunswick, NJ

Bachelor of Science

Rutgers University
Computer Engineering

Certificates

10/2021

Google Cloud Professional Data Engineer

Google Cloud