

Sanket Thakre

Seattle, WA | 657-253-9649 | sanket.thakre3@gmail.com | linkedin.com/in/srthakre3 | github.com/srthakre3 | sanketthakre.me

ANALYTICS ENGINEER • DATA ENGINEER

Business Intelligence Engineer at Amazon, five years across pharmacy and fulfillment operations. I build and own ETL pipelines, dimensional models and the data quality systems underneath them. Strongest on inherited systems: two of my main projects arrived unowned and on fire, and both are still in production.

EXPERIENCE

Business Intelligence Engineer *Amazon Pharmacy, Pharmacy Operations Data* Jun 2025 - Present

- Migrated the full performance guarantee reporting stack for two pharmacy benefit manager partners off another org's infrastructure: 12 contractual metrics, 46 ETL pipelines and 24 warehouse tables across dev and prod, plus SFTP delivery to external partners.
- Repeatedly catch reporting logic errors on contractual metrics before they trigger financial penalties. Audited one after a threshold breach, proved 40 fulfillment errors had been wrongly counted against Amazon, and drove the fix that corrected the quarter.
- Resolved a Sev-2 contract compliance defect in partner call metrics, correcting the calculation from 100% of calls to the contractual top 80%, then deploying and backfilling history so partner-facing dashboards were accurate.
- Built the first proactive data quality alerting on metrics carrying financial exposure, where none existed before, and cleared 7 governance violations that would otherwise have disabled contractual reporting jobs.
- Automated the weekly operations speed business review and the daily executive flash report, both previously assembled by hand, delivering a PDF and formula-level Excel workbook to a self-service subscriber list.
- Launched three QuickSight dashboards on fulfillment learning curves and pharmacy services productivity, used as evidence in State Board of Pharmacy inspections and regulatory accreditation.

Business Intelligence Engineer *Amazon, Central Operations (IDEA / COBI, later QuBIT)* Aug 2021 - Jun 2025

- Took over the model behind Amazon's global fulfillment operational planning weeks before the 2025 annual cycle, when the maintaining team moved orgs. Reverse-engineered nine interconnected stored procedures from a short handover and delivered the cycle on schedule across NA, EU, Japan and India.
- Primary developer and product owner for the North America operations alerting platform serving 100+ sites, used by fulfillment leaders to manage backlog and customer experience risk. Wrote the original architecture and design document as my first project at Amazon, built the real-time alerting alongside another engineer, then transitioned it to a new owner.
- Took over an unowned pipeline feeding shift-level productivity data available in no other source system, with two director escalations open. Resolved the data quality and missing data errors behind them and kept it running; its dashboard saw 30-40K monthly views spanning operations, program, business and director-level users. Still in production.
- Became the reference point for two BIEs and three business analysts across four regions: wrote the onboarding documentation, ran architecture training, and mentored on stored procedure optimization and automated data validation.
- Ran Redshift SQL tuning and tooling sessions, wrote SOPs for metric development and on-call intake, maintained runbooks for Peak and Prime, and designed the team's database access request process, which became the team standard.

TECHNICAL SKILLS

Languages SQL (expert), Python (Pandas, NumPy, PySpark), Bash

Data ETL/ELT design, pipeline migration, dimensional modeling (Kimball, star schema, SCD), data quality engineering, monitoring and alerting

Cloud & Warehouses Amazon Redshift, AWS (Athena, S3, Glue, Lambda, IAM), PostgreSQL, MySQL

BI Amazon QuickSight, Tableau

Practice Git, code review, dev/prod separation, GitHub Actions, runbooks and SOPs, on-call

PERSONAL PROJECTS

End-to-End ELT Pipeline *dbt · Airflow · PostgreSQL · Docker · GitHub Actions* 9.5M rows of NYC TLC data through staging, intermediate and dimensional mart layers (star schema); 22 passing dbt tests, Airflow DAGs with alerting callbacks, documented lineage.

Data Quality Framework *Python · Pandas · SQLAlchemy · pytest* Enforces data contracts on warehouse tables: schema validation, null and duplicate detection, referential integrity, statistical outlier alerting, CI quality gates.

Self-Serve Analytics Platform, E-Commerce *SQL · dbt · Redshift · QuickSight* Raw e-commerce data through a dbt dimensional model into a governed metric layer (revenue, churn, cohort retention) with documented lineage and tested business logic, so non-technical stakeholders self-serve.

EDUCATION

M.S. Computer Engineering California State University, Fullerton

2021

B.E. Electronics & Telecommunications University of Pune, India

2018