

Sanket Thakre

Analytics Engineer

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

TECHNICAL SKILLS

Data Engineering: ETL/ELT Pipelines, dbt, Apache Airflow, Fivetran, Airbyte, Dimensional Modeling, Star/Snowflake Schema, SCD Types, Data Quality

Languages: SQL (Advanced), Python (Pandas, NumPy, PySpark, Matplotlib), Bash

Cloud & Warehouses: Amazon Redshift, Snowflake, Databricks, AWS S3, AWS Glue, MySQL, PostgreSQL, Oracle

BI & Visualization: Amazon QuickSight, Tableau, Power BI, Looker, Jupyter Notebook

Tools: Git, GitHub Actions, Agile/Scrum, Data Contracts, Performance Tuning, Root Cause Analysis

EXPERIENCE

Amazon · Seattle, WA

Aug 2021 – Present

Business Intelligence Engineer

- Designed and implemented scalable ETL/ELT data pipelines using SQL and Python to ingest, transform, and load large datasets into Amazon Redshift, serving as the data backbone for multiple product and operations teams.
- Developed and maintained dimensional data models — star/snowflake schemas and SCD types — to standardize KPI definitions across cross-functional teams, enabling consistent, reliable self-serve analytics.
- Automated end-to-end data collection, transformation, and delivery workflows, eliminating manual reporting overhead and enabling real-time insight delivery through dashboards, deep-dive tools, and alerting systems.
- Built and maintained production dashboards and metrics tooling in Amazon QuickSight, providing real-time operational visibility to business, finance, and operations stakeholders.
- Partnered cross-functionally with data scientists, software development engineers, product managers, and business analysts to gather requirements, define data contracts, and ship end-to-end analytics solutions.
- Performed complex deep-dive analysis on key business metrics to identify root causes, surface actionable insights, and drive high-impact strategic decisions and process improvements.

PROJECTS

End-to-End ELT Pipeline — dbt & Apache Airflow

Python · dbt · Apache Airflow · PostgreSQL · GitHub Actions

- Ingesting NYC TLC trip data via Python into PostgreSQL, transforming with dbt (staging → intermediate → dimensional mart with star schema), orchestrating with Airflow DAGs, and enforcing data quality via dbt tests and CI pipeline.

Automated Data Quality & Monitoring Framework

Python · Great Expectations · Pandas · SQL

- Reusable Python framework for automated schema validation, null/duplicate detection, and statistical outlier alerting on Redshift/PostgreSQL tables — simulating production-grade data observability for analytics pipelines.

Self-Serve Analytics Platform — E-commerce Metrics

SQL · dbt · Redshift · Tableau / QuickSight

- End-to-end analytics solution: raw e-commerce data → dbt dimensional model → governed metric layer → interactive dashboard (revenue, churn, cohort retention) with documented data lineage and design rationale.

EDUCATION

Master of Science, Computer Engineering

California State University, Fullerton · Fullerton, CA

May 2021

Bachelor of Engineering, Electronics & Telecommunications

University of Pune · Pune, India

May 2018