

Sharad Sharma

+91-880-044-6941 | s13sharad@gmail.com | Gurgaon, India

[linkedin.com/in/sharadsharma12](https://www.linkedin.com/in/sharadsharma12)

SUMMARY

Senior Data Engineer with **9+ years** building large-scale data platforms. Currently at **Amazon (Fintech)**, I own tax data pipelines handling **45B+ records/month** across US and Global markets. Strong background in **PySpark, Airflow, AWS (EMR/Redshift/S3)**, dimensional modeling, and data warehouse design. Previously built Hadoop-based platforms at Optum and TCS.

SKILLS

Python, SQL, Java, Scala, Shell Scripting | PySpark, Spark, Hadoop, Hive, Sqoop

AWS: EMR, S3, Redshift, Glue, Athena, Lambda, CloudWatch, QuickSight, Lake Formation, CDK, IAM

Apache Airflow, DBT, Oozie | Dimensional Modeling, SCD, Data Warehouse Design

PostgreSQL, Oracle, HBase | Git, CI/CD

EXPERIENCE

Data Engineer II - Amazon India (Fintech), Gurugram

Apr 2022 -

Present

Own tax data platform for AWS-Global (**15B+ records/month, \$12B+ revenue, 15 seller-of-record countries**) and US-Retail (**30B+ records/month, \$45B+ revenue, ~1B records/day during peak**).

Designed **Bronze to Silver to Gold pipeline architecture** with backfill support and data completeness scoring; built the full stack independently from ingestion through reporting.

Manage **34 Airflow DAGs** total - 18 for Global (incl. 8 recon DAGs with business rules) and 16 for US-Retail. Revenue/tax reconciliation across OFA (AP, AR, GL), ATARs, and TAR.

Migrated legacy reporting to native AWS (Airflow + EMR + Redshift). **Improved monthly SLA from Day 8 to Day 5, cut infra costs 71%**.

Built archival framework **reducing Redshift query latency 54%**. 9 QuickSight dashboards + real-time ETL monitoring. **Saves compliance teams ~1,200 hours/year**.

Security: Secrets Manager rotation, least-privilege IAM, Lake Formation controls, CloudTrail auditing.

Senior Data Engineer - To The New, Noida

Jan 2022 - Apr

2022

Led client onboarding for auction trade-analysis platform (Airflow, DBT, Spark SQL). Managed requirements, documentation, and **6-month delivery roadmap single-handedly**.

Data Engineer - Optum / UnitedHealth Group, Noida

Aug 2020 - Jan

2022

Ran **50+ scheduled pipelines** ingesting **5+ TB/day** from DB2, Oracle, SQL Server into Hadoop (**100TB cluster, 750+ nodes**). Built monitoring framework with automated QC reports and alerting.

Created end-to-end automation for analytics/data-science workflows - **reduced manual effort ~80%**.

Software Data Engineer - Tata Consultancy Services, Thane

Sep 2016 - Apr

2020

Built Hadoop analytics platform (Hive, Sqoop, MapReduce, custom UDFs). **Led 5-person team**.

Developed full-stack monitoring dashboard (Angular, D3.js, Java, PostgreSQL) - **adopted org-wide** for production deployment tracking. **Won Applause Award**.

EDUCATION

B.Tech, Computer Science - G.L. Bajaj Institute of Technology and Management, Greater Noida (2016)

CERTIFICATIONS

Six Sigma Green Belt | Developer of the Month - Optum/UHG