



Shubham Vinodkumar Yadav

SENIOR CONSULTANT

✉ Shubham.yadav.5497@gmail.com 📠 7021722812 📍 Mumbai, India 🌐 <https://linkedin.com/in/shubham-yadav-000a75167>



Summary

Senior Data Engineer with 7+ years of experience building and supporting data pipelines on **Azure Databricks** and **SAS-based enterprise data platforms**. Good hands-on experience with **Databricks, PySpark, Delta Lake, Delta Live Tables (DLT), Medallion Architecture, SCD-Type 1, SCD Type 2, CI/CD** and **Spark SQL**, along with extensive experience in **SAS DI, SAS Viya, and SAS programming**. Experienced in designing **scalable batch and streaming pipelines**, optimizing Spark jobs, data modeling and collaborating with clients to deliver **production-ready data pipeline**



Projects

Actuarial Modernization SAS To GCP Migration

Senior Consultant Jun 2026 - Present

- Migrating actuarial data assets and ETL workloads from **SAS** to Google **Cloud Platform (GCP)**.
- Assessing SAS jobs, datasets, libraries, and dependencies to support migration planning.
- Developing data **export jobs** to generate **compressed Parquet** files and load data into Google Cloud Storage (GCS).
- Converting SAS programs into PySpark for execution on Google Cloud Dataproc.
- Validating migrated data by **reconciling record counts, schemas, and business outputs** between SAS and GCP.
- Collaborating with client, infrastructure, and security teams to ensure successful migration and environment readiness.

Modern Data Ecosystem Build - Austrailian Cleint

Senior Consultant Aug 2025 - May 2026

- Designed and developed **scalable ETL pipelines** using **Azure Databricks, PySpark, and Delta Lake**, supporting both **batch and streaming workloads**.
- Implemented **Medallion Architecture (Bronze, Silver, Gold)** to improve data quality, reprocessing, and downstream analytics performance.
- Built **Delta Live Tables (DLT)** pipelines with data quality expectations and automated dependency management.
- Optimized Spark jobs by applying **partitioning, broadcast joins, and shuffle reduction**, improving pipeline performance.
- Collaborated directly with **clients and stakeholders** to gather requirements, design solutions, and deliver production ready pipelines.

Snowflake Data Engineering – POC Project

Lead Developer Mar 2025 - Jul 2025

- Built an **end-to-end ETL pipeline** in Snowflake using **Medallion Architecture**.
- Ingested data from **Amazon S3, Azure Blob Storage, and Azure SQL DB**.
- Developed **Snowpipe, Streams, Tasks, and Stored Procedures** for automated data flow. Set up a **Snowflake Clean Room** to enable secure data sharing and collaboration.

Campaign Management System and Automation - Malaysia Client

Senior Developer Nov 2023 - Mar 2025

- Developed data pipelines to load data from source systems into campaign databases.
- Automated ETL pipelines and monitoring processes for improved reliability.
- Developed and executed C1360 campaigns targeting specific customer segments.



Experience

Senior Consultant

Datacurate Technology Private LTD

May 2023 - Present

Consultant

Datacurate Technology Private LTD

Apr 2022 - Apr 2023

Associate Consultant

Datacurate Technology Private LTD

Aug 2019 - Mar 2022



Skills

- Databricks, PySpark, Spark SQL, Delta Lake, Delta Live Tables, Structured Streaming, ADF, GCP Data Proc
- Azure (ADLS, Blob Storage, Azure SQL)
- SQL, Python
- Base SAS, Advance SAS, SAS DI/SAS ETL, SAS/ODS, SAS Viya 3.5
- Snowflake, Postgres, MySQL, Oracle



Languages

English

Hindi



Education

Bachelor of Engineering

Xavier Institute of Engineering

Mar 2015 - Mar 2019

Post Graduate Diploma in Data Engineering

IIT-Jodhpur

Feb 2024 - Jun 2026



Certification

Snowflake (SnowPro Core)

Azure DP-203

Azure Fundamentals

Data Warehousing and Reporting

Senior Developer *Feb 2020 - Oct 2023*

- Designed and maintained **enterprise ETL pipelines** using **SAS DI and SAS Viya** for data warehousing and reporting.
- Optimized reporting workflows, reducing processing time from **4+ hours to 20 minutes**.
- Integrated SAS applications with multiple client systems for seamless data exchange.
- Built SAS web services to enable smooth data flow between SAS and client platforms.
- Designed business management dashboards using SAS Viya for enhanced data visualization.