

Muhammad Abubakar

Senior Data Engineer

📍 Lahore, Pakistan | ✉ abubakar.dataengineer@gmail.com | ☎ +92 335 4631993

🌐 LinkedIn: [muhammadabubakar12/](https://www.linkedin.com/in/muhammadabubakar12/)

Professional Summary

Results-driven **Senior Data Engineer** with **4+ years** of experience building scalable cloud-based data pipelines, ETL/ELT workflows, and data warehouses on **AWS**. Expertise in **PySpark, AWS Glue, EMR, Athena, Step Functions, and Redshift**, with a strong background in **SQL/NoSQL databases (PostgreSQL, DynamoDB)** and **real-time data processing**. Proven ability to optimize data infrastructure, enforce governance, and collaborate with cross-functional teams to deliver actionable insights.

Technical Skills

- **Cloud Technologies:** AWS EMR, Glue, Athena, Step Functions, S3, Redshift, RDS, DynamoDB, CloudWatch, AWS CLI
- **Big Data & ETL:** PySpark, Databricks, Oracle Data Integrator (ODI), Medallion Architecture
- **Databases:** PostgreSQL, MySQL, Oracle, MongoDB, DynamoDB, Delta Lake
- **Programming:** Python (Data Processing), SQL, PL/SQL, Shell Scripting
- **Data Governance:** Security, Encryption, Compliance (TBAML, KYC, FCCM)
- **Tools:** Stonebranch, Postman, Toad, Oracle WebLogic

Work Experience

Senior Data Engineer

Techlogix | Lahore, Pakistan

Sep 2022 – Present

- Designed and deployed **scalable ETL pipelines** using **AWS S3, Databricks, and PySpark**, reducing Databricks costs by **25%** through Delta Lake optimizations.

- Architected **data warehousing solutions** on AWS, integrating structured (PostgreSQL) and unstructured (DynamoDB) data sources for analytics.
- Automated complex workflows using **shell scripting** and **Stonebranch**, improving operational efficiency by **30%**.
- Managed **petabyte-scale healthcare data** for Blue Cross Blue Shield (BCBS), ensuring compliance with **AWS security best practices**.
- Collaborated with compliance teams to implement **TBAML/KYC frameworks**, leveraging **ODI 12c** and **SQL Server**.

Software Engineer

Hypoters | Remote

Oct 2020 – Nov 2021

- Developed **RESTful APIs** using Node.js and MongoDB, enabling real-time data access for international clients.
- Optimized **MySQL databases** for performance and scalability.

Key Projects

AWS-Based Data Pipeline Optimization

Techlogix | 2023–2024

- Built **PySpark-powered ETL pipelines** on **AWS EMR** to process large-scale financial data, improving throughput by **40%**.
- Leveraged **AWS Glue** for metadata management and **Athena** for ad-hoc querying.
- Implemented **AWS Step Functions** to orchestrate multi-stage workflows, reducing manual intervention.

Cloud Data Warehouse for BCBS

Techlogix | 2022–2023

- Architected petabyte-scale data pipeline processing 50M+ daily records using **PySpark on AWS EMR**, reducing costs by 25% via Delta Lake optimizations
- Implemented **Medallion Architecture** on S3 (Bronze→Gold) with AWS Glue Catalog, enabling self-service analytics via Athena

- Automated 30+ workflows using **Step Functions** and Lambda, cutting manual efforts by 40%
- Enforced HIPAA compliance through **IAM policies**, KMS encryption, and CloudWatch monitoring

Real-Time Trade Monitoring (TMX Repository)

Techlogix | 2023–Present

- Designed data models and ETL architecture for trade-related red flags in the TMX repository.
- Built transformation pipelines to populate OFSDW with normalized trade data.
- Integrated CBS APIs for data updates via RESTful services.
- Led production rollout and provided ongoing L2 support and SLA compliance.

FCCM Implementation & Customization

Techlogix | Feb 2022 – May 2023

- Built custom ETL workflows in ODI 12c to support TBAML scenarios in line with SBP regulations.
- Defined data mappings in collaboration with compliance teams for AML detection.
- Developed DIS files for ingestion by the FCCM engine.
- Performed validation, testing, and integration using SQL Server, Postman, and Toad.
- Implemented shell scripting to automate batch executions, improving operational efficiency.
-

Education

Bachelor of Science in Computer Science (BSCS)

National University of Sciences and Technology (NUST) | *2017–2021*

Achievements

- 🏆 **Achiever of the Month** – Techlogix (*Sep 2022, Jun 2023, Jul 2024*)