Kamal Hasan Ansari

📍 Delhi, UP | 📧 kamalhasan.cse.19@email.com | 📞 (807) 754 3240 | LinkedIn: [www.linkedin.com/in/kamal-hasan-ansari-021842248](http://www.linkedin.com/in/kamal-hasan-ansari-021842248) | GitHub: https://github.com/tech-kamal

# Professional Summary

Detail-oriented Data Engineer with 1 year of experience in designing ETL pipelines, data modeling, and optimizing workflows in cloud and on-prem environments. Skilled in SQL, Python, Spark, Airflow, and AWS, with strong exposure to building scalable data pipelines, data quality checks, and reporting automation. Adept at collaborating with cross-functional teams to deliver reliable, production-grade data solutions.

# Technical Skills

• Programming & Scripting: Python, SQL  
• ETL & Orchestration: Apache Airflow, dbt, Talend  
• Big Data & Cloud: PySpark, AWS (S3, Redshift, EC2, Lambda), Azure Data Factory  
• Databases: PostgreSQL, MySQL  
• Data Visualization: Tableau, Power BI  
• Version Control & DevOps: Git, Docker, CI/CD

# Professional Experience

## Data Engineer (Intern/Entry-Level)

Hi Tech Brains Solutions PVT LMT – Bareilly | Jan 2025 – Aug 2025

• Designed and implemented ETL pipelines in Airflow and Python to ingest and process structured/unstructured data from APIs, databases, and flat files.  
• Optimized SQL queries and Spark jobs, improving data processing efficiency by 30%.  
• Automated data validation scripts to ensure compliance with quality and governance standards.  
• Deployed data pipelines on AWS (S3, Redshift, Lambda) reducing manual reporting efforts.  
• Collaborated with analysts and stakeholders to design Tableau/Power BI dashboards, enabling real-time insights.  
  
Key Achievement: Reduced reporting latency from 6 hours to under 1 hour by implementing optimized scheduling and parallelized ETL workflows.

# Projects

• Real-Time Data Pipeline with Kafka & Spark  
 Built a real-time streaming pipeline using Kafka, Spark Streaming, and AWS S3 to process clickstream data. Applied aggregations and pushed processed data to Redshift for BI reporting.  
  
• Healthcare Claims Data Warehouse  
 Developed ETL workflows in Airflow + dbt to transform raw claims data into a Snowflake data warehouse. Created star schema models to support regulatory compliance and fraud detection dashboards.

# Education

Bachelor of Technology in Computer Science  
Dr A P J Abdul Kamaal Technology Lucknow – 2023

# Certifications

• AWS Certified Data Engineer – Associate  
• Databricks Lakehouse Fundamentals  
• Microsoft Azure Data Engineer Associate