

# SHAHEER JAMAL

Karachi, Pakistan

+92 323 8178936 | [Email](#) | [LinkedIn](#) | [Github](#) | [Portfolio](#) | [Linktree](#)

## OBJECTIVE

Motivated Computer Science student specializing in Cloud Data Engineering with hands-on experience building scalable ETL pipelines and real-time data processing systems. Proficient in orchestrating workflows using Apache Airflow and architecting event-driven solutions on AWS (Glue, Lambda, Athena). Skilled in Python, SQL, and Snowflake for data warehousing and automation. Eager to leverage a strong technical foundation in cloud infrastructure and big data technologies to solve complex data challenges in a professional setting

## EXPERIENCE

### Associate Cloud Data Engineer Intern | Saylani Mass Training Programme (Feb 25 – Dec 25)

- Architected and deployed enterprise-grade AWS ETL pipelines and Airflow orchestration workflows.
- Served as a Teaching Assistant, mentoring students in Cloud Data Engineering, Python, and advanced SQL

### Software Developer | Al Quran Academy (April 24 – Jan 25)

- Engineered an end-to-end Academy Management Software to automate daily administrative and student tracking operations

## PROJECTS

### Pandemic Insights: ETL Pipeline with AWS Glue & Athena | [Github](#)

- Architected a scalable ETL pipeline to ingest and transform complex COVID-19 datasets from disparate sources into a centralized Data Warehouse.
- Engineered optimized Data Models (Star Schema) in Amazon Redshift, designing fact and dimension tables to reduce query latency by 40%.
- Orchestrated serverless data transformations using AWS Glue jobs and enabled ad-hoc SQL analysis via Amazon Athena for immediate insights generation

### Automated Weather Data Pipeline with Apache Airflow | [Github](#)

- Orchestrated an end-to-end ETL workflow using Apache Airflow to automate the extraction of real-time weather data from the OpenWeather API.
- Designed Python-based DAGs (Directed Acyclic Graphs) to manage task dependencies, ensuring fault-tolerant scheduling and efficient data flow control.
- Implemented automated data ingestion into AWS S3, transforming raw JSON API responses into structured formats for historical analysis and reporting

### Event-Driven Data Processing System (AWS) | [Github](#)

- Designed a decoupled, event-driven architecture leveraging AWS SNS, SQS, and Lambda to handle high-throughput asynchronous data streams.
- Implemented robust message queuing patterns to ensure data durability and fault tolerance during traffic spikes.
- Optimized Lambda execution logic to minimize cold starts and achieve real-time processing triggers for dynamic event payloads

### Real-Time Stock Market Data Pipeline | [Github](#)

- Built a high-velocity streaming pipeline using Apache Airflow and AWS to capture and process real-time stock market fluctuations.
- Integrated AWS Glue Crawlers and Athena to automate schema discovery and enable near real-time SQL querying of streaming data stored in S3.
- Developed a fault-tolerant ingestion layer ensuring sub-second data availability for downstream analytics dashboards.

### Real-Time Data Pipeline with SCD (Snowflake & NiFi) | [Github](#)

- Developed a resilient data ingestion workflow using Apache NiFi and Python to stream synthetic data into Amazon S3.
- Implemented Slowly Changing Dimensions (SCD Type 1 & 2) logic to accurately track historical data changes and maintain dimension history.
- Automated data loading into Snowflake using Snowpipe, ensuring continuous, low-latency updates for the data warehousing layer.

## SKILLS

- Cloud Platforms (AWS):** S3, Redshift, Glue, Athena, Lambda, DynamoDB, SNS, SQS, EC2, IAM.
- Data Engineering & Streaming:** Snowflake (Snowpipe), Apache Airflow, Apache Kafka, Apache NiFi.
- Core Concepts:** Python, Advanced SQL, ETL/ELT Pipelines, Data Modeling (Star Schema), Slowly Changing Dimensions (SCD), EDA.

## EDUCATION

### GRADUATION

BSCS | Sir Syed University of Engineering & Technology, 2028

### HIGH SCHOOL

Computer Science | Delhi Science College, 2024

### SECONDARY SCHOOL

Computer Science | TFS Schooling System, 2022

## CERTIFICATIONS

### Cloud Data Engineering | Nov 2025 (SMIT)

Comprehensive training covering AWS, Python, ETL pipelines, and Data Warehousing.

### Snowflake Hands-On Essentials | Snowflake Inc.

Demonstrated practical proficiency in building pipelines, schemas in snowflake

### Associate Data Engineer | (in progress)

Currently mastering advanced ETL patterns and data lakehouse architecture.