

Rohan Kulkarni

Senior Cloud Data Engineer

rohan.k@email.com — +91-8XXXXXXXXXX — LinkedIn — GitHub

Professional Summary

Senior Cloud Data Engineer with 7+ years of experience architecting, building, and operating large-scale cloud-native data platforms. Specialized in AWS-based data lakes, batch ETL frameworks, and analytics-ready data layers. Strong background in Python-driven data processing, SQL optimization, and workflow orchestration using Apache Airflow. Proven track record of owning data systems end-to-end, from ingestion to BI consumption, while optimizing cost, reliability, and performance.

Core Technical Skills

- **Languages:** Python, SQL
- **Cloud Platform:** AWS (S3, Glue, Redshift, Lambda, EC2, IAM)
- **Data Engineering:** Batch ETL, ELT, Data Lakes
- **Orchestration:** Apache Airflow
- **Warehousing:** Redshift, Dimensional Modeling
- **Data Quality:** Validation, Reconciliation, Monitoring
- **DevOps:** CI/CD, Git, CloudWatch

Professional Experience

CloudNova Solutions

Senior Cloud Data Engineer

Bangalore, India

Jan 2020 – Present

- Architected a cloud-native enterprise data lake on Amazon S3 ingesting data from 20+ internal and external systems
- Designed multi-layer data architecture (raw, curated, analytics) enabling analytics and machine learning workloads
- Built scalable ETL pipelines using AWS Glue and Python processing **1.5TB+ data daily**
- Led migration of on-premise data warehouse to Amazon Redshift with zero critical downtime
- Implemented table partitioning, compression, and lifecycle policies reducing storage and query costs by **30%**
- Orchestrated complex workflows using Apache Airflow with dependency management, retries, and SLA monitoring
- Developed data validation and reconciliation frameworks to ensure accuracy across ingestion layers
- Partnered with BI and analytics teams to design analytics-ready datasets and KPI tables
- Acted as technical owner for production data pipelines, handling incidents and root-cause analysis

InnovaTech Systems

Data Engineer

Pune, India

Jul 2017 – Dec 2019

- Developed Python- and SQL-based ETL pipelines for ingesting data from RDBMS and REST APIs
- Automated batch workflows using Airflow, improving reliability and observability
- Supported reporting and analytics use cases by delivering clean, modeled datasets
- Assisted in legacy system migration projects to cloud-based data platforms

Selected Projects

AWS Enterprise Data Lake Platform

- Designed scalable data lake architecture supporting analytics and ML workloads
- Enabled self-service analytics for analysts and data scientists
- Standardized ingestion patterns across teams

Cost Optimization & Monitoring Framework

- Implemented CloudWatch-based monitoring for data pipelines
- Optimized compute and storage usage across AWS services
- Reduced monthly infrastructure spend significantly

Education

Bachelor of Engineering in Information Technology

XYZ University

2017

Certifications

- AWS Certified Data Analytics – Specialty
- AWS Certified Solutions Architect – Associate