Sandeep Rao

Lead Data Engineer
sandeep.rao@email.com — +91-8XXXXXXXXX — LinkedIn — GitHub

Professional Summary

Lead Data Engineer with 9+ years of experience designing, scaling, and governing enterprise data platforms in cloud environments. Proven expertise in building high-volume batch ETL systems, defining data architecture standards, and leading engineering teams responsible for mission-critical analytics pipelines. Strong focus on reliability, data quality, cost optimization, and aligning data platforms with long-term business strategy.

Core Technical Skills

- Languages: Python, SQL
- Cloud: AWS (S3, Glue, Redshift, Lambda, IAM)
- Orchestration: Apache Airflow
- Data Architecture: Data Lakes, Warehouses
- Modeling: Star & Snowflake Schemas
- Governance: Data Quality, Access Control
- DevOps: CI/CD, Monitoring, Alerting

Professional Experience

InsightWorks Technologies Hyderabad, India
Lead Data Engineer Feb 2020 – Present

- Architected and owned an enterprise AWS-based data lake ingesting **2TB+ data daily** from product, finance, and operations systems
- Defined multi-layer architecture (raw, refined, analytics) supporting BI dashboards and data science workloads
- Led a team of 6 data engineers, establishing coding standards, review processes, and on-call rotations
- Designed standardized ETL frameworks in Python and AWS Glue, improving pipeline consistency across teams
- Implemented Airflow-based orchestration with SLA monitoring, retries, and alerting
- Improved pipeline success rate from 92% to **99.8%** by addressing failure patterns and performance bottlenecks
- Introduced data quality checks, reconciliation logic, and anomaly detection for critical datasets
- Partnered with analytics, finance, and leadership teams to define KPIs and data delivery SLAs
- Played key role in data roadmap planning, cost forecasting, and platform scalability decisions

TechNova Solutions Chennai, India
Senior Data Engineer Jun 2016 – Jan 2020

- Built batch ETL pipelines using Python and SQL for analytics and reporting use cases
- Supported migration from on-prem data warehouse to cloud-based platforms
- Optimized SQL queries and data models to improve reporting performance

Key Initiatives

Enterprise Data Platform Modernization

- Migrated legacy reporting systems to cloud-native architecture
- Enabled self-service analytics for business users
- Reduced infrastructure and operational overhead significantly

Education

Bachelor of Engineering — Computer Science 2016