# FULL NAME

City, State | Email: email@example.com | Phone: +91-XXXXXXXXXX | LinkedIn: linkedin.com/in/username | GitHub: github.com/username

## Professional Summary

Data Engineer with expertise in building scalable data pipelines, optimizing ETL workflows, and implementing cloud-native big data solutions. Proficient in SQL, Python, Spark, and distributed systems with hands-on experience in AWS and Azure. Strong background in data modeling, real-time analytics, and delivering high-performance solutions for business insights.

## Core Technical Skills

Programming: Python, SQL, Java  
Big Data: Apache Spark, Hadoop, Hive, Kafka  
Databases: PostgreSQL, MySQL, MongoDB, Snowflake  
Cloud & DevOps: AWS (S3, Redshift, EMR, Glue), Azure Data Factory, Docker, Airflow  
Data Engineering: ETL, Data Warehousing, Data Modeling, Streaming Pipelines  
Tools: Git, Jenkins, Linux, CI/CD, Agile/Scrum

## Professional Experience

**Data Engineer – Company Name**  
City, State | MM/YYYY – Present

• Designed and deployed automated ETL pipelines using Airflow, Python, and AWS Glue, reducing data processing time by 40%.  
• Built real-time streaming pipelines with Kafka and Spark Structured Streaming to process high-volume data with <5s latency.  
• Optimized Redshift queries and partitioning strategies, improving query performance by 35%.  
• Collaborated with analysts and ML engineers to deliver clean, reliable datasets for advanced analytics and machine learning.

**Data Engineer Intern – Company Name**  
City, State | MM/YYYY – MM/YYYY

• Assisted in developing Spark-based batch jobs to migrate legacy ETL workflows to the cloud.  
• Automated data quality checks using Python, SQL, and AWS Lambda, improving accuracy by 25%.  
• Documented pipeline workflows and implemented version control with Git and Jenkins.

## Projects

• Data Lake Architecture – Designed and deployed a cloud-based data lake on AWS S3 with Glue catalog and Athena for analytics.  
• Real-Time Log Analytics – Built a Kafka-Spark pipeline to monitor and process server logs, enabling proactive issue detection.  
• Cloud Cost Optimization – Automated cost-tracking dashboards using Python and Redshift, saving ~15% in cloud spend.

## Education

B.Tech in Computer Science | XYZ University, City | YYYY

## Certifications

• AWS Certified Data Analytics – Specialty  
• Microsoft Azure Data Engineer Associate  
• Google Cloud Professional Data Engineer