Data Engineer – Job Description

About the Role:

We are looking for a highly skilled Data Engineer to join our delivery team. The role involves building robust, scalable, and secure data pipelines that power analytics and reporting. The ideal candidate will have strong experience in cloud platforms, ETL frameworks, and big data tools, and will work closely with analysts, business teams, and clients to ensure the timely delivery of clean, high-quality data for strategic insights and decision-making.

Key Responsibilities:

✓ Design and build scalable ETL/ELT pipelines for structured and unstructured data.

✓ Develop data infrastructure using cloud services (AWS Glue, S3, Redshift, etc.).

✓ Optimize data storage, processing, and query performance using PySpark or similar tools.

✓ Implement and maintain data security, governance, and access control mechanisms.

✓ Monitor and troubleshoot pipeline performance to ensure high availability and reliability.

✓ Collaborate with business and analytics teams to understand and meet data requirements.

✓ Ensure seamless integration of data from internal and external sources.

✓ Create clear technical documentation for pipelines, workflows, and architecture.

✓ Participate in client meetings, provide status updates, and highlight risks or blockers.

✓ Drive automation, standardization, and process improvements across data engineering functions.

Skills & Qualifications:

• **Programming**: Strong in SQL, Python, or Scala

• **Big Data**: Experience with PySpark, Spark, or Hadoop

• **Cloud Platforms**: Hands-on with AWS, Azure, or GCP (Glue, S3, Redshift, BigQuery, etc.)

• **ETL Tools**: Familiarity with Airflow, Azure Data Factory, or other orchestration tools

• **DevOps & CI/CD**: Experience with Git, pipeline automation, and deployment tools

• **Security & Governance**: Knowledge of data compliance and best practices

• **Monitoring Tools**: Experience with CloudWatch, Prometheus, or Datadog

• **Education**: Bachelor's degree in computer science, Engineering, or related field

• **Certifications**: AWS/Azure/GCP cloud certifications are a plus

Experience:

✓ 2–5 years of experience in Data Engineering or similar roles

✓ Strong experience in building cloud-based data pipelines and real-time data processing