**Job Summary:**

As Senior Associate L2 in Data Engineering, you will translate client requirements into technical design, and implement components for data engineering solution. Utilize deep understanding of data integration and big data design principles in creating custom solutions or implementing package solutions. You will independently drive design discussions to insure the necessary health of the overall solution 

The role requires a hands-on technologist who has strong programming background like Java / Scala / Python, should have experience in Data Ingestion, Integration and data Wrangling, Computation, Analytics pipelines and exposure to Hadoop ecosystem components. You are also required to have hands-on knowledge on at least one of AWS, GCP, Azure cloud platforms.

**Role & Responsibilities:**

Your role is focused on Design, Development and delivery of solutions involving:

·         Data Integration, Processing & Governance

·         Data Storage and Computation Frameworks, Performance Optimizations

·         Analytics & Visualizations

·         Infrastructure & Cloud Computing

·         Data Management Platforms

·         Implement scalable architectural models for data processing and storage

·         Build functionality for data ingestion from multiple heterogeneous sources in batch & real-time mode

·         Build functionality for data analytics, search and aggregation

**Experience Guidelines:**

***Mandatory Experience and Competencies:***

|  |  |
| --- | --- |
| **#** | **Competency** |
| 1. | Overall 5+ years of IT experience with 3+ years in Data related technologies |
| 2. | Minimum 2.5 years of experience in Big Data technologies and working exposure in at least one cloud platform on related data services (AWS / Azure / GCP) |
| 3. | Hands-on experience with the Hadoop stack – HDFS, sqoop, kafka, Pulsar, NiFi, Spark, Spark Streaming, Flink, Storm, hive, oozie, airflow and other components required in building end to end data pipeline. |
| 4. | Strong experience in at least of the programming language Java, Scala, Python. Java preferable |
| 5. | Hands-on working knowledge of NoSQL and MPP data platforms like Hbase, MongoDb, Cassandra, AWS Redshift, Azure SQLDW, GCP BigQuery etc |
| 6. | **Well-versed and working knowledge with data platform related services on at least 1 cloud platform, IAM and data security** |

***Preferred Experience and Knowledge (Good to Have):***

|  |  |
| --- | --- |
| **#** | **Competency** |
| 1. | Good knowledge of traditional ETL tools (Informatica, Talend, etc) and database technologies (Oracle, MySQL, SQL Server, Postgres) with hands on experience |
| 2. | Knowledge on data governance processes (security, lineage, catalog) and tools like Collibra, Alation etc |
| 3. | Knowledge on distributed messaging frameworks like ActiveMQ / RabbiMQ / Solace, search & indexing and Micro services architectures |
| 4. | Performance tuning and optimization of data pipelines |
| 5. | CI/CD – Infra provisioning on cloud, auto build & deployment pipelines, code quality |
| 6. | **Cloud data specialty and other related Big data technology certifications** |

**Personal Attributes:**

·         Strong written and verbal communication skills

·         Articulation skills

·         Good team player

·         Self-starter who requires minimal oversight

·         Ability to prioritize and manage multiple tasks

·         Process orientation and the ability to define and set up processes