**Job Profile: Data Engineer**

**What you’ll do**

* Work on Big Data engineering projects using Spark (PySpark primarily) and Databricks
* Perform analytics on existing data using data frames, complex joins, optimizing the code, and migration processes (Strong SQL is required)
* Work on migration projects, including database migration from general pool app, warehouses Redshift database migration, or Informatica data to database migration
* Work on orchestration on various tools such as Data Factory, AWS Glue, Informatica, etc.
* Have a good understanding of working knowledge of SQL
* Understand ETL and orchestrate via ETL data
* Work on DevOps and know how to deploy code
* Have fundamental knowledge of PySpark, Scala Spark, or Java Spark
* Understand Cloud solutions and deploy code and spin up clusters on Data bricks or Azure

**Must Have Skills**

* 1-2 years of experience in Spark (PySpark) and Azure DE technologies (primarily Databricks and ADF)
* Strong SQL knowledge, especially complex querying, query optimizations, stored procedures, triggers, window functions, partitioning, etc.
* Knowledge of PySpark and/or fundamental knowledge of Scala Spark or Java Spark
* Experience in working with Data Bricks is a must have
* Knowledge of SQL data warehousing concepts or Azure Synapse
* Experience working with ETL
* Familiarity with Azure DE ecosystem, especially ADB and ADF
* Working knowledge of Git and GitHub
* Strong and clear foundations in the basics

**Good to Have Skills**

* Knowledge of AWS DE tools such as AWS Glue, AWS EMR, AWS Redshift, etc.
* Proficiency in ETL tools such as Informatica
* PL/SQL, T-SQL, etc.
* Snowflake, or Teradata is a good to have
* Working knowledge of DevOps
* Experience working with GCP DE ecosystem

**Communication Skills**

* Candidate should have excellent communication skills
* Candidate should have the drive to get the work done. They should have high accountability and take ownership of the work assigned to them and get it done at any cost
* Ability to communicate unambiguously to the stakeholders and delegate tasks to subordinates
* Ability to represent technical concepts and thoughts verbally in clear terms
* Ability to give scenario-based examples, use case and applications of technical concepts