Data Engineer

Job Summary:

We are seeking a skilled Big Data Engineer with expertise in Python, Hadoop, Spark, Shell scripting, PostgreSQL, and Linux to join our data engineering team. The ideal candidate will be responsible for designing, developing, and maintaining scalable data pipelines and platforms to support our data analytics and business intelligence needs.

About iSteer Technologies:

iSteer's mantra of **"taking the quantum leap"** with "smarter technology" solutions, is producing what customers wanted in a fraction of time, with fewer people, less overall cost and getting the product to market when it is still relevant to target customers. We truly believe "technology must not only support our customer's business, but it must advance it at a faster pace". At iSteer (ISO 27001:2013 certified), we enable our customers, achieve a competitive advantage with technology solutions and products that are derived from decades of cross-industry experience and technology expertise. iSteer helps transform the business of several global enterprises with their technology needs across industries.

We provide **Digital Integration, Data Engineering and connect enterprises by providing Robotic Process Automation, IOT, Cloud, and AI solutions**, through our world-class product engineering expertise, our products like AppSteer make it easier to transform businesses digitally. We have exponentially grown our operations across the globe with 250+ employees **at our offices in India, Singapore, United States, Canada and Dubai**. Our expansion globally has always been a remarkable difference and delivers key results to our customers. Our renowned partners are Workato Platinum Partner, TIBCO Gold Partner and Dell Boomi. Life at iSteer, where a fine line of young and experienced minds leads into the infinite opportunities in the digital era.

At iSteer, we make sure that talent meets technology in a culture which is driven by knowledge and growth. Being a part of iSteer makes you a stakeholder of achievements which will turn your latent potential into a success story. It is also enabled by excellence into our culture which encourages individual development, embraces an inclusive environment, rewards innovative excellence and supports our communities.

Why join us
- Competitive salary + benefits
- We value work-life balance and encourage taking vacations
- Excellent leadership and an outstanding network of individuals
- On-going regular training certifications sponsored by iSteer.

Key Responsibilities:

- Develop, maintain, and optimize large-scale data processing workflows using **Hadoop** and **Apache Spark**.
- Write efficient, reusable, and maintainable code in **Python** and **Shell scripting** for data ingestion, transformation, and automation tasks.
- Design and manage **PostgreSQL** databases including schema design, query optimization, and performance tuning.
- Manage Linux-based environments and tools for data platform deployment, monitoring, and troubleshooting.
- Collaborate with data scientists, analysts, and other stakeholders to understand data requirements and deliver reliable data solutions.
- Implement data quality checks and data validation processes.
- Monitor and optimize data pipeline performance and resource utilization.
- Document workflows, processes, and system architecture to ensure maintainability and knowledge sharing.

Required Skills and Qualifications:

- Strong programming skills in **Python** for data processing and automation.
- Hands-on experience with **Hadoop ecosystem components** (HDFS, MapReduce, YARN).
- Proficiency in **Apache Spark** for distributed data processing and analytics.
- Expertise in **Shell scripting** (Bash, etc.) for task automation and system management.
- Experience working with **PostgreSQL** database systems, including advanced SQL and performance tuning.
- Strong familiarity with **Linux** operating systems and command-line tools.
- Understanding of data modeling, ETL processes, and big data architecture.
- Familiarity with version control systems such as Git.
- Excellent problem-solving skills and ability to work in a collaborative team environment.
- Good communication skills, both written and verbal.

Preferred Qualifications:

- Experience with containerization (Docker, Kubernetes) and cloud platforms (AWS, Azure, GCP).
- Knowledge of other big data tools like Kafka, Hive, or Airflow.
- Prior experience working in Agile/Scrum teams.