



Job Description

Sr. Data Engineer

You are interested in large-scale data analysis, aggregation, data processing, and storage. We expect you to work closely with others and provide leadership, code discipline, and project design by example. You are extremely driven and want to assume ownership over an important part of our technology to build a solid product stack. You will use your expert Java skills to design and build high-load web applications and service-oriented systems for storing, processing, and searching a very large volume of unstructured text. In doing so, you will solve a wide variety of engineering challenges, ranging from data flow, to storage, to aggregation, to supporting APIs / presentation layer / data pipelines / web apps. You are comfortable with change. You've had positive experience working for a startup before.

Required Skills

- 4+ years of Python or Java or Scala using Spark Framework
- Expert SQL knowledge of one of the following: MySQL, PostgreSQL, Oracle, SQL Server
- Strong shell/bash scripting skills
- Experience with:
 - Designing, building, and maintaining data processing systems
 - Schema design and dimensional data modeling
 - Running production-grade systems
 - Spark Framework
 - ClickHouse
- Bonus = working experience with:
 - Big data stack components: Kafka, Nifi, Elastic, Scylla DB, Manticore Search
 - Distributed system concepts and big data technology stacks
 - Data science/analysis
 - Kubernetes
 - NoSQL

Responsibilities

- Collaborate with analytics and business teams to improve data models that feed business intelligence tools
- Increase data accessibility and foster data-driven decision making across the organization
- Define company data assets (data models), Spark / Spark SQL to populate data models
- Design data integrations and data quality frameworks
- Design and evaluate open-source and vendor tools for data lineage