

Technology Division

Data & Artificial Intelligence Department

Position: Senior Data Engineer

Job Description:

- Design, build, and maintain robust **data pipeline architectures** to support data integration, analytics, and business intelligence.
- Lead the development of complex ETL/ELT pipelines to extract, transform, and load data from diverse sources (databases, APIs, flat files, Excel, XML, CSV) ensuring high availability and data integrity.
- Develop and manage large, complex datasets that meet functional and non-functional business requirements.
- Identify, design, and implement process improvements including **automation** of manual tasks, **optimization of data delivery**, and **infrastructure re-design for scalability**.
- Identify and eliminate manual bottlenecks through automation, infrastructure re-design, and modern CI/CD practices.
- Implement and support **data integration tools** on Azure and Pentaho for enterprise-scale data workflows.
- Collaborate with **BI Analysts, Data Scientists, and Product teams** to ensure pipelines provide actionable insights on customer acquisition, operational efficiency, and business KPIs.
- Develop data tools and infrastructure that empower analytics teams and support **AI/ML initiatives**.
- Enhance and maintain database performance, reviewing logs, system metrics, and user reports.
- Partner with Database Administrators, Developers, and Analysts to improve data management, governance, and security practices.
- Drive engineering excellence by conducting code reviews, establishing documentation standards, and mentoring junior team members.
- Implement advanced data security, row-level permissions, and disaster recovery protocols in compliance with financial industry standards.
- Proactively monitor and troubleshoot complex batch/streaming jobs, optimizing SQL queries and system metrics to ensure sub-second data availability.

Preferred Skills:

- **Advanced SQL & Database Mastery:** Expert-level SQL (Stored Procedures, Indexing, Functions, Triggers, Views, Indexing) and deep experience with **PostgreSQL, Oracle, and MSSQL**.
- Hands-on experience with **PostgreSQL, Oracle, and MSSQL** in both on-premises and **cloud environments (Azure SQL, Azure Synapse, Data Lake)**.
- Practical knowledge of **Pentaho Data Integration (Kettle), or Apache Hops** and **Azure data integration tools** (ADF, Data Factory pipelines, Data Bricks is a plus).
- Experience with messaging technologies such as **Kafka** or **Azure Event Hub**.
- Experience building and optimizing **big data pipelines, architectures, and data sets**.
- Strong knowledge of **data warehouse architecture, ETL processes, and metadata management**.
- Strong analytical skills with the ability to work with **structured and unstructured datasets**.
- Proven ability to troubleshoot data pipeline and database performance issues.
- Experience performing **root cause analysis** on internal and external data flows.
- Strong knowledge of **data security, permissions, recovery, and monitoring**.
- Strong proficiency in **Python** or **PySpark** for advanced data transformation and workflow automation.
- Excellent **communication skills** and ability to collaborate with cross-functional teams (Executives, Product, Data, Design).
- **Financial services industry knowledge** is preferred.
- Strong organizational skills, detail-oriented, and able to work under pressure with tight deadlines.
- Positive team player with a **growth mindset** and willingness to learn and share knowledge.