**Role Description & Person Profile**

|  |  |
| --- | --- |
| **Job Title** | Senior Data Engineer |
| **Business** | Central |
| **Location** | Head Office, Peterborough |
| **Job Level** | 11 |
| **Team Structure**Reports to and direct reports | Reports to Data Engineering Lead |
| **Budget Responsibility** | N/A |

|  |  |
| --- | --- |
| **Role Overview**Impact Statement  | We are looking for a Data Engineer with proven experience to join our growing team of analytics experts. As the senior Data Engineer will be responsible for expanding and optimizing our data and data pipeline architecture, as well as optimizing data flow and collection for cross functional teams. The ideal candidate is an experienced data pipeline builder and data wrangler who enjoys optimizing data systems and building them from the ground up. You will support our software engineers, database architects, data analysts and data scientists on data initiatives and will ensure optimal data delivery architecture is consistent throughout ongoing projects. You must be self-directed and comfortable supporting the data needs of multiple teams, systems and products. The right candidate will be excited by the prospect of optimizing or even re-designing our company’s data architecture to support our next generation of products and data initiatives. |
| **Key Responsibilities** | We are looking for a candidate with proven experience (**at least two projects using Databricks / Fabric for modern data processing**) in a Data Engineer role, who has attained a Graduate degree in Computer Science, Statistics, Informatics, Information Systems or another quantitative field. * Create and maintain optimal data pipeline architecture,
* Assemble large, complex data sets that meet functional / non-functional business requirements.
* Identify, design, and implement internal process improvements: automating manual processes, optimizing data delivery, re-designing infrastructure for greater scalability, etc.
* Build the infrastructure required for optimal extraction, transformation, and loading of data from a wide variety of data sources using SQL/Non-SQL and Azure ‘big data’ technologies.
* Build analytics tools that utilize the data pipeline to provide actionable insights into customer acquisition, operational efficiency and other key business performance metrics.
* Work with stakeholders including the Executive, Product, Data and Design teams to assist with data-related technical issues and support their data infrastructure needs.
* Keep our data separated and secure across national boundaries through multiple data centres and Azure regions.
* Create data tools for analytics and data scientist team members that assist them in building and optimizing our product into an innovative industry leader.
* Work with data and analytics experts to strive for greater functionality in our data systems.
* For developing, constructing, maintenance and testing of data architectures, such as databases and large-scale processing systems
* Clean, aggregate, and organize data from disparate sources and transfer it to data storage for analysis. Making sure the companies data technology is operating at its peak results in massive improvements to performance, cost, or both.
* Owning the stability of new products designed, including the on-going robustness, resilience and stability of these products;
* Building and maintaining custom ingestion pipelines,
 |
| **Key Stakeholders** | The IS Development Team is responsible for multiple enterprise systems used by employees and customers on a global scale. **Internal:*** BSM / Product Development teams / End users – collaborate on solution developments depending on delivery methodology
* Technical consultants – Collaborate on defining and developing technical solutions and propose practical technical recommendations.
* Technology professions – proactive support for Technology professions

You will work closely with:* Cross functional within central D&A team
* Business Systems Managers
* Business Systems Advisors
* IS Project Managers
* Key and Specialist Users
 |
| **Other Factors**Travel, shift pattern, working hours, Licence type etc. | 37.5 hrs per week, permanent |

|  |
| --- |
| **Required experience, qualifications, and necessary knowledge**  |
| **Essential** | **Desirable** |
| * Excellent strategic thought process.
* Advanced working SQL knowledge and experience working with relational databases, query authoring (SQL) as well as working familiarity with a variety of databases.
* Experience building and optimizing data pipelines, architectures and data sets.
* Experience performing root cause analysis on internal and external data and processes to answer specific business questions and identify opportunities for improvement.
* Strong analytic skills related to working with unstructured datasets.
* Build processes supporting data transformation, data structures, metadata, dependency and workload management.
* A successful history of manipulating, processing and extracting value from large, disconnected datasets.
* Experience supporting and working with cross-functional teams in a dynamic environment.
* Databricks Spark 3.0 Certified
* Databricks Certified Data Engineer Professional.
* Azure Data Engineer Associate Certified
* Experience with relational SQL and NoSQL databases
* Experience with Data Integration
* Experience with Relational Databases and Data Warehousing

**Key Skills:*** Fabric / Azure Data Factory
* Data Lake Storage (gen 2)
* Databricks
* Lakehouse Architecture
 | * Working knowledge of message queuing, stream processing, and highly scalable ‘big data’ data stores.
* Strong project management and organizational skills.
* Fabric Associate Certified
* Experience with stream-processing systems: Spark-Streaming, Azure IOT / Event Hub, Azure Stream Analytics, Kafka
* Experience with object-oriented/object function scripting languages: Python (E), Java
* Experience with Data Archiving
* Experience with MS SQL, Logic Apps, Function Apps, SSIS, SSRS (D)Experience with Microsoft Azure DevOps (D)Experience with Data Visualisation (D)
* Experience with Cosmos DB, Azure Analysis Services, Synapse Analytics (SQL & Spark Pools)
 |

|  |  |
| --- | --- |
| **Key Behaviours** |  |
| **AB Agri High Performance Framework** | Our high-performance framework is a set of guiding behaviours which have been created with people from across our businesses to enable great performance across the organisation. The focus is on what you can do to demonstrate high performance in your role, as well as the behavioural inputs to assist you getting there. * Pioneering – Curious, spirited and bold. We lead the right way.
* Excellence – We seek excellence in all that we do.
* Growth – We create ways for our people and customers to thrive. That’s how we keep making a difference.
 |