**Kishore M**

[Kishoretex003@gmail.com](mailto:Kishoretex003@gmail.com) | +1 (469) 215-5007 | [linkedin.com/in/kishore-m-823040294](https://www.linkedin.com/in/kishore-m-823040294)hadoop

**PROFESSIONAL SUMMARY:**

* With 8+ years of experience in data engineering, Tableau, BI, and data analytics make me an exceptionally talented and experienced Senior Azure Data Engineer.
* Capable of creating, deploying, and refining data engineering solutions on the Azure platform; proficient in Tableau and Power BI analytics and reporting.
* Scalability and resilience for Azure Websites, as well as diagnostics, monitoring, and analytics on Azure Platform, were all implemented together with the deployment and configuration of Azure Websites.
* Developed Azure Web roles, deployed Azure SQL Services, and managed Azure Storage.
* Simplified database organization and data extraction using SQL and NoSQL queries from MySQL, PostgreSQL, and SQL Server databases.
* Controlling and providing database access while migrating SQL databases to Azure cloud services likes Azure Data Lake, Azure Data Lake Analytics, Azure SQL Database, Data Bricks, and Azure SQL Data warehouse utilizing Azure Data Factory to migrate on-premises databases to Azure Data Lake stores.
* Knowledgeable about utilizing Tableau and Power BI for comprehensive reporting, data visualization, and data wrangling.
* Proficient in T-SQL development, able to create intricate searches by Combining several tables, and has exceptional skills in creating and managing stored procedures, triggers, and user-defined functions.
* Expert in creating and putting into practice Azure-based data solutions, including data processing, analysis, and storage using Azure cloud services.
* Knowledgeable about developing and putting into use scalable and dependable data pipelines with PySpark, Azure SQL, ADLS, Databricks, and Azure Data Factory (ADF).
* Seasoned expert with knowledge of Kafka, with a focus on real-time data integration and stream processing.
* Strong understanding of RDBMS like SQL Server and NoSQL databases like MongoDB, as well as strong writing and optimization skills for SQL queries utilizing functions, packages, and triggers.
* Worked together with cross-functional teams to collect business needs and convert them into data solutions that are technological.
* I have developed ETL solutions in Azure Databricks utilizing Spark SQL for data Extraction, Transformation, and aggregation from various file formats and data sources. The data is then analyzed and transformed to reveal customer usage trends.
* Proficient in pipeline building, including the creation of connected servers, datasets, and dynamic code in Azure Data Factory.
* Moving from the current ETL system to Azure Data Factory.
* Possess expertise in developing data models, data flows, and Power BI dashboards and reports.
* Configuring Databricks with Microsoft Azure, Using Databricks Workspace for Business Analytics, and Managing Clusters in Databricks.
* Used Hadoop scripts for data loading and manipulation on the HDFS (Hadoop File System).
* Utilizing the Spark execution engine, accumulated vast and intricate data on the Apache Hadoop platform.
* DevOps tools such as GitHub and Azure DevOps have been used in the development of CI/CD pipelines for application migration to various environments.
* I used confluence to document all my work while working on JIRA for defect/issue tracking and reporting.
* Positiveness, openness to learning innovative ideas, and readiness to take on difficulties characterize this initiative-taker.

**TECHNICAL SKILLS:**

|  |  |
| --- | --- |
| **Languages** | Python, PySpark, SQL, T-SQL, Unix, Shell Scripting |
| **Frameworks** | Spark, Spark Structured streaming, Kafka, Hadoop |
| **Cloud** **Technologies** | Azure, Databricks, Kafka, Snowflake |
| **Azure** | Azure Data Factory, Azure Data Lake Storage Gen2, Azure SQL DB and SQL DW (Azure Synapse), Azure DevOps, Power Apps, Azure Key Vault, Logic Apps |
| **SQL and NoSQL Databases** | MongoDB, MySQL, Teradata, Oracle, SQL server |
| **Visualization Tools** | Power BI, Tableau |
| **Source Control** | GitHub, Azure DevOps |
| **File Extensions** | CSV, JSON, XML, Excel, DAT |
| **Utilities** | JIRA, Confluence, ServiceNow |
| **Tools** | UiPath, Power Automate, Airflow |
| **SDLC** | Agile Methodology, Waterfall Methodologies |
| **Soft Skills** | Leadership, Team Management, Scrum Management |

**WORK EXPERIENCE:**

**Company: Humana October 2022 to Present**

**Role: Senior Azure Data Engineer Louisville, Kentucky, US**

**Responsibilities:**

* Established a centralized Data Lake with Databricks and Azure Data Factory, tailored for efficient processing of medical data in the healthcare domain.
* Engineered robust Azure Data Factory pipelines to automate data preparation, ensuring accuracy and compliance with healthcare regulations.
* Applied SQL, PySpark, and Azure Databricks for medical data processing, meeting stringent privacy and security requirements in healthcare.
* Orchestrated ETL procedures with Azure Data Lake Analytics, seamlessly integrating diverse medical data sources into Azure-based healthcare ecosystems.
* Architected scalable Data Fabric solutions on Azure, enhancing data accessibility and privacy compliance for healthcare analytics initiatives.
* Implemented CI/CD pipelines in Azure DevOps for secure deployment of medical data pipelines across healthcare environments.
* Managed high-performance Data Lake environments within Azure Databricks, crucial for rapid and compliant processing of medical data.
* Executed streamlined ETL plans to consolidate medical data into Azure Synapse Analytics, ensuring data integrity and compliance with healthcare standards.
* Leveraged Agile methodologies and GIT for efficient project management within healthcare data engineering teams.
* Conducted rigorous testing on medical reports to optimize performance and ensure compliance with healthcare regulations.
* Utilized multi-dimensional models like Star and Snowflake schema to design healthcare data warehouses tailored to medical analytics needs.
* Demonstrated expertise in migrating medical databases to Snowflake, enhancing scalability and compliance for healthcare analytics.
* Designed optimized fact and dimension tables to facilitate advanced medical data analysis.
* Fine-tuned API Gateway configurations to optimize performance and ensure data privacy across healthcare applications.
* Developed logical models to drive insights and inform decision-making across healthcare data engineering initiatives.
* Engaged in various healthcare data engineering facets, including batch processing and real-time message ingestion, to drive innovation and compliance in healthcare analytics.

**Environment**: Azure Data Factory, Azure Databases, Azure Data Lake Storage Gen2, Azure Data Fabric, Azure Blob Storage, SQL, Python, PySpark, GIT, Azure DevOps, Snowflake, Confluence, JIRA

**Company: Availity July 2019 to September 2022**

**Role: Senior Data Engineer Jacksonville, Florida, US**

**Responsibilities:**

* Established efficient data intake pipelines for seamless transfer of medical data to the Data Lake from diverse sources.
* Automated data preparation processes with Azure Databricks, reducing modeling time and effort for medical data.
* Developed Spark applications on Azure Databricks to aggregate medical data, enabling comprehensive analysis of client usage patterns.
* Implemented delta tables and automated operations in Databricks for accurate and timely extraction of medical data.
* Deployed the Medallion Architecture for structured medical data storage, facilitating quick access and analysis.
* Oversaw data processing workflows to ensure consistency and dependability in healthcare data management.
* Demonstrated proficiency in managing Big Data platforms and relational databases for healthcare data.
* Designed large-scale data solutions on Snowflake Cloud Data Warehouse for real-time collaboration.
* Integrated TensorFlow and Scikit-learn into Azure Machine Learning for improved model accuracy.
* Configured cloud deployments and automated builds for healthcare Java applications using CI/CD pipelines.
* Developed Azure Functions for event-driven automation, enabling resource efficiency in healthcare solutions.
* Ensured data security and compliance by integrating Azure Key Vault within the Data Fabric architecture.
* Leveraged ADF and Apache Spark to process medical data efficiently.
* Created ADF pipelines to transform and load medical data from various sources, enhancing data exchange capabilities.

**Environment:** Azure Data Factory, Azure Data Fabric, Azure Data Lake Storage Gen 2, Azure Blob Storage, ADF, Azure, SQL, Python, PySpark, GIT Azure DevOps, MongoDB, Snowflake, Confluence, JIRA, ServiceNow

**Company: Decathlon May 2018 to June 2019**

**Role: Azure Data Engineer Hyderabad, Telangana, India**

**Responsibilities:**

* Oversaw the establishment of efficient data collection channels, ensuring seamless integration of insights from various platforms into the Data Lake.
* Improved analysis by automating and refining data preparation processes with Azure Databricks, optimizing comprehension of preferences and trends.
* Developed Spark applications on Azure Databricks for extracting and examining data, revealing valuable insights into buying patterns and market trends.
* Implemented advanced data models to securely handle data, ensuring its precision and reliability.
* Managed data pipelines and ETL operations to smoothly convert data, improving its usefulness for targeted marketing and interactions.
* Leveraged Amazon's expertise to supply scalable cloud-based storage solutions, enabling flexible and scalable management of data.
* Demonstrated proficiency in SQL and Hive for efficient processing of datasets, facilitating rapid analysis of trends.
* Enhanced data security by applying encryption and tokenization in the API Gateway, safeguarding sensitive data throughout its lifecycle.
* Applied Waterfall and Agile methodologies to ensure timely delivery of insights, fostering adaptability.
* Emphasized code-health and testing practices to streamline data processing, ensuring precise insights.
* Enabled the shift to cloud-based processing with Azure Data Lake Store and Azure Data Factory, empowering with advanced analytics capabilities.
* Collaborated cross-functionally to create comprehensive pipelines for integrating and analyzing data, enabling informed decision-making and enriching experiences.

**Environment:** MongoDB, SQL, Python, Azure Data Fabric, Azure cloud services, GIT, DevOps, SQL Auto loader, AI/ML Algorithms, ServiceNow, Confluence, JIRA

**Company: L&T Finance June 2015 - May 2018**

**Role: ETL Developer Hyderabad, Telangana, India**

**Responsibilities:**

* Leveraged PySpark for data extraction, cleaning, transformation, and loading into a Hive data warehouse, ensuring financial data accuracy and consistency.
* Utilized Spark to create RDDs for efficient data processing from various sources, enhancing performance and scalability in finance analytics.
* Developed Spark jobs with window functions for corporate reporting and analytics, ensuring accurate financial insights and compliance.
* Ensured data correctness and consistency in finance ETL processes through SSIS-managed extraction, transformation, and loading activities.
* Demonstrated proficiency in Hive partitioning and bucketing principles for optimized financial data storage and retrieval.
* Implemented encryption and tokenization in the API Gateway to enhance financial data security during transmission and storage.
* Engaged in the full SDLC lifecycle for finance big data projects, ensuring comprehensive requirement analysis, design, coding, testing, and production deployment.
* Actively pursued knowledge in open-source big data technologies like Hadoop and NoSQL databases to enhance financial data management capabilities.
* Possessed strong understanding of SDLC phases and business analysis abilities, crucial for successful finance data projects.
* Demonstrated competence in finance data management areas, including Data Lake, Business Intelligence, and Data Warehousing, focusing on ETL and SDLC processes.
* Created Azure data pipelines with airflow operators for finance ETL tasks, ensuring efficient data processing and analysis.
* Implemented sophisticated ETL functionalities in Spark or Pandas data frames for enhanced finance data manipulation and analytics.
* Executed daily query jobs to load financial data from Google Cloud storage into SQL tables, ensuring timely and accurate financial reporting.
* Responsible for creating, testing, and maintaining ETL procedures in finance data management, ensuring data integrity and reliability.
* Conducted unit testing and optimization of ETL code and SQL queries to improve finance data processing performance and efficiency.
* Possessed experience in SQL databases with large transaction volume, ensuring robust finance data management in production environments.
* Developed PL/SQL stored procedures and SQL scripts for managing finance data in MySQL databases, enhancing data integrity and reliability.
* Proficient in R and Python for finance data manipulation, statistical analysis, and modeling, enabling informed financial decision-making and forecasting.

**Environment:** MySQL, ETL, Azure, Teradata, SSIS, JIRA, Oracle, Hive, Spark, Python, Spark, Agile

**Education**: Bachelor of Electronics and Communication Engineering at Gandhi Institute of Technology and Management, Visakhapatnam, AP, India. Graduation Date: 2015