Supriya Doppalapudi

Data Engineer

📞 +1 9727462519 🌐 supriyad976@gmail.com ⚲ United States 🌐 linkedin.com/in/supriya-doppalapudi

SUMMARY

Experienced Data Engineer with over 4 years of expertise in crafting and optimizing data solutions. I specialize in making data work seamlessly. Skilled in data warehousing and big data technologies and have good experience in ETL tools ensuring data flows smoothly. My focus is on enhancing data governance, security, and compliance standards. I take pride in implementing innovative technologies like Databricks and Snowflake, making data processing faster and more efficient.

TECHNICAL SKILLS

- **Programming Languages:** Python, Java, Scala
- **Data Storage:** SQL, NoSQL databases, Amazon Redshift, Google BigQuery, Snowflake, Data Lakes
- **Big Data Technologies:** Apache Hadoop, Apache Spark, Databricks, Apache Flink
- **Data Processing Frameworks:** Apache Kafka, Apache NiFi
- **ETL Tools:** Apache Airflow, Informatica
- **Cloud Platforms:** Amazon Web Services, Microsoft Azure, Google Cloud Platform
- **Containerization and Orchestration:** Docker, Kubernetes
- **Data Integration:** Apache Camel, MuleSoft
- **Version Control:** Git
- **Database Management Systems:** MySQL, PostgreSQL, Oracle, Snowflake
- **Data Quality and Profiling:** Apache Griffin, Trifacta, Talend Data Quality
- **Collaboration Tools:** Jira, Confluence
- **Visualization tools:** Tableau, Power Bi
- **Scripting and Automation:** Shell scripting, PowerShell
- **Continuous Integration/Continuous Deployment:** Jenkins, GitLab CI
- **Data Governance:** Apache Atlas, Collibra
- Machine Learning

PROFESSIONAL EXPERIENCE

Delta Airlines   January 2023 – Present
Data Engineer

- Spearheaded the design and implementation of robust data pipelines, optimizing data flow efficiency and enabling faster and more accurate data processing.
- Orchestrated the management of databases, ensuring seamless integration and enhancing data quality, contributing to improved overall system reliability.
- Innovated and implemented advanced technologies such as Databricks and Snowflake, driving increased data processing speed and enabling real-time analytics for business stakeholders.
- Collaborated with cross-functional teams to develop and enforce security and compliance standards, ensuring enhanced data governance and mitigating potential risks.
- Researched and recommended new technologies, standard processes, and tools, leading to the adoption of scalable solutions that improved overall data infrastructure efficiency.
- Built robust systems to collect, manage, and convert raw data into usable information, resulting in increased data accessibility for data scientists and business analysts.
- Implemented ETL processes that streamlined data transformation, reducing processing time and enhancing the accuracy of data interpretation for analytics teams.
- Played a key role in the successful migration to a data lake architecture, resulting in reduced storage costs and improved scalability for future business needs.
- Leveraged expertise in non-relational databases, contributing to improved handling of unstructured data and facilitating better utilization for analytical purposes.
- Executed projects with a focus on scalability and efficiency, demonstrating a commitment to optimizing resource utilization in data management and processing.

Tata Consultancy Services (State Farm)   December 2020 – July 2022
Data Engineer

- Pioneered the development and implementation of robust ETL processes using Databricks, significantly enhancing overall data accuracy and efficiency.
- Collaborated closely with the analytics team to design and optimize Snowflake data warehouses, elevating data accessibility and seamlessly integrating with diverse reporting tools.
- Engineered scalable solutions for data lakes, leveraging advanced Python scripts to automate data ingestion, cleaning, and transformation.
- Established and maintained a streamlined data governance framework, ensuring compliance with industry standards and reducing data errors through rigorous validation processes.
- Led the migration of legacy systems to modern cloud-based platforms, achieving cost reductions and improved scalability for future data needs.
- Developed a real-time data monitoring system using Apache Kafka, enabling proactive identification and resolution of data anomalies.
- Orchestrated the seamless integration of scalar databases into the existing data infrastructure, optimizing data retrieval times and providing the analytics team with expedited access to critical business insights.
- Mentored junior team members in utilizing advanced data engineering concepts and tools, resulting in improved team efficiency and project delivery timelines.
- Collaborated with cross-functional teams to define and implement data quality metrics, contributing to heightened overall data reliability.
- Spearheaded the automation of end-to-end testing for ETL pipelines, reducing manual effort and fortifying the overall reliability of data processing workflows.

Tata Consultancy Services (PepsiCo)   April 2019 – November 2020
Data Engineer - Intern

- Engineered and executed robust ETL pipelines using Python, Airflow, Spark, and SQL, enhancing data processing efficiency for diverse source systems.
- Architected data models and implemented warehousing solutions on major cloud platforms AWS, Azure, Snowflake, reducing data retrieval times.
- Executed data extraction, transformation, and cleaning processes with meticulous attention, ensuring a high level of data accuracy and elevating overall data quality.
- Automated critical data processes, optimizing workflow efficiency and scalability, resulting in increased team productivity for analytics and machine learning initiatives.
- Collaborated closely with analysts and scientists to decipher complex data requirements, fostering a collaborative environment that accelerated project delivery timelines.
- Monitored and maintained data infrastructure, implementing proactive measures to ensure consistently high performance and reliability.
- Spearheaded the adoption of best practices in data engineering, contributing to team professional development and facilitating knowledge-sharing sessions.
- Identified and rectified bottlenecks in data pipelines, improving real-time data accessibility for analytical and product development efforts.
- Implemented a comprehensive data governance framework, contributing to a reduction in data anomalies and enhancing overall data integrity and security.
- Established key performance indicators and conducted regular data quality audits, supporting informed decision-making.

CERTIFICATIONS

Microsoft Azure Fundamentals: The Microsoft Azure Fundamentals certification validates foundational knowledge of Azure cloud concepts and services. It exhibits expertise with Azure pricing, security, core services, and infrastructure. By obtaining this certification, I can demonstrate my proficiency in using Azure to create safe and scalable solutions, efficiently managing Azure resources, and comprehending Azure pricing

structures. This certification on demonstrates proficiency with Azure basics and makes a valuable candidate for Azure-related positions and opportunities.

AWS Certified Cloud Practitioner (In progress)**:** This certification attests to your fundamental understanding of AWS services and cloud computing. It exhibits mastery of key AWS services, security and compliance, pricing models, architectural best practices, and cloud concepts. Obtaining this certification allows to demonstrate the capacity to build scalable and secure solutions by adhering to industry best practices, optimizing costs, and making effective use of AWS services. This certification will demonstrate proficiency with AWS and cloud computing, which will lead to a variety of cloud-related jobs and career opportunities.