Vanamala Bhargav

Data Engineer

Location: New Mexico | **Phone:** +1 (505) 490-6084 | **Email:** bhargavvanmala@gmail.com

PROFESSIONAL SUMMARY

Highly skilled and analytical Data Engineer with over 4 years of experience in designing, developing, and optimizing complex data systems within the IT sector. Proficient in deploying scalable cloud solutions, advanced data pipelines, and real-time data processing frameworks using a variety of tools and technologies. Known for enhancing data-driven decision-making processes, improving operational efficiencies, and delivering actionable insights. Adept at leading projects that streamline operations, ensuring data integrity and security, and driving strategic business growth and innovation. Seeking to leverage expertise in a dynamic role that supports strategic business objectives and fosters innovation.

TECHNICAL SKILLS

Programming Languages: Python, SQL, Java, C#
Databases: SSIS, Apache Spark, SQL Server, Oracle XE, Apache Hive, Hadoop, AWS Glue
Data Visualization: ETL, Tableau, PowerBI, MS Excel
Cloud Technologies: AWS Services (EC2, S3, Redshift, EMR, AWS Glue)
Tools & Technologies: Git, Docker, PySpark, Apache Kafka
Others: Advanced SQL, Common Table Expressions (CTEs), subqueries, window functions, joins, CSV, Parquet TXT, Data Governance, Compliance Frameworks, Data Encryption, Project Management

EDUCATION

Master of Science in Computer Science

University of Dayton, Dayton, OH, United States   January 2022

Bachelor of Engineering in Computer Engineering

Jawaharlal Nehru Technological University, Hyderabad, India   July 2019

WORK EXPERIENCE

Data Engineer | Deloitte, New Mexico   April 2023 – Present
- Optimized data retrieval by designing efficient stored procedures, functions, triggers, and views across critical databases.
- Uncovered actionable insights by composing complex SQL queries using CTEs, subqueries, window functions, and joins on large datasets.
- Maintained data integrity with Python scripts for AWS Glue Catalog metadata access and repair queries.
- Ensured data consistency and accuracy by architecting Java Spark ETL frameworks for master data transformation.
- Enhanced source data quality through advanced techniques like partitioning, deduplication, and type casting.
- Ensured compliance by securing sensitive data transfer to Spark jobs in EMR Clusters using AWS Parameter Store.
- Improved data accessibility by strategizing file storage in CSV, Parquet, and TXT formats using S3.
- Enhanced data quality by developing C# console applications for preprocessing and validating incoming data.
- Boosted SQL performance through rewriting queries, updating statistics, and adding indexes.

Data Engineer | Citizen's Bank, New Mexico   January 2022 – April 2023
- Enhanced data availability and decision-making by leading the implementation of real-time data processing solutions with Apache Kafka and Spark.
- Optimized enterprise-level data storage and retrieval through robust data warehousing solutions on AWS Redshift.
- Reduced manual intervention and error rates by automating data pipeline workflows using Python and Apache Airflow.
- Facilitated predictive analytics and business intelligence by collaborating with data scientists to deploy machine learning models into production.
- Ensured regulatory compliance by implementing comprehensive data governance frameworks adhering to GDPR and other standards.
- Safeguarded sensitive information through enhanced data security measures, including encryption and fine-grained access controls.

Data Engineer | Jukshio Technology Innovation Private Limited, India   August 2019 – July 2021
- Ensured smooth data flow and reliability by designing, developing, and maintaining data pipelines and ETL processes using SQL Server Integration Services (SSIS).
- Enhanced data quality by developing custom C# scripts within SSIS to preprocess and cleanse raw data, addressing header mismatches and empty datasets.
- Facilitated complex data transformations by creating intricate SQL queries, stored procedures, and functions across diverse data sources.
- Improved query performance with advanced SQL techniques, including window functions, common table expressions and

relevance (CTEs), and recursive queries.
- Streamlined automated daily processing by managing SSIS package deployment and configuring the SSIS DB catalog and schedules.
- Optimized data integration solutions through collaboration with cross-functional teams to gather requirements and design effective strategies.
- Reduced SQL query execution times by up to 90% through performance tuning, including indexing, rewriting, and analysis.
- Adapted to evolving business needs by refactoring legacy SQL queries and redesigning reports for improved performance and relevance.

Core Qualifications

- **Cloud Data Architectures:** Proficient in designing and implementing scalable, secure cloud data solutions using AWS technologies, including Redshift, S3, EMR, and Glue.
- **Data Pipeline Development:** Skilled in developing robust data pipelines that ensure efficient data flow, processing, and storage, utilizing tools like Apache Spark, Apache Kafka, and SSIS.
- **Real-Time Data Processing:** Expert in setting up real-time data streams and processing frameworks to enhance operational decision-making and analytics capabilities.
- **Data Security and Compliance:** In-depth knowledge of implementing data security measures and compliance protocols, ensuring adherence to GDPR and other regulatory standards.
- **Performance Optimization:** Proven track record in optimizing data access and processing speeds through advanced SQL tuning, indexing strategies, and ETL process enhancements.
- **Machine Learning Deployment:** Experienced in collaborating with data scientists to integrate machine learning models into production environments, boosting predictive capabilities and business insights.