

Santhoshbharadwaj H A

Data Engineer

<https://www.linkedin.com/in/santhoshbharadwaj-h-a-2b009615b/> | Phone: +91 8970210996 | santhoshbharadwaj2498@gmail.com

EDUCATION:

BE. in Information Science

Visvesvaraya Technological University

2017-2021

Diploma in computer Computer Application

Shri Swami Vivekananda institute

Graduated: June 2017

Technical SKILLS:

Programing Languages:

Python, javascript, Bash, shell Scripting, scala

Big Data Technologies:

PySpark, Apache Spark

Databases: SQL

Cloud Platform: Amazon Web Services(AWS)

Data Processing And

Analysis:

Pandas

ETL Tool: Apache Airflow, AWS Glue

Version Control: Git

IDE: Jupyter Notebook, Pycharm, Vs Code

AWARDS:

Rotary Youth Award (2021)

Chess: Qualified for state level

LANGUAGES:

Technical language (Proficiently)

English (Proficiently)

Kannada (Intermediate)

Hindi (Intermediate)

Telugu (Intermediate)

PROFESSIONAL SUMMARY:

- Data Engineer with over 4 years of experience specializing in designing and implementing scalable data solutions, with deep expertise in Python, PySpark, SQL, AWS, and Pandas.
- Led the development and maintenance of large-scale ETL pipelines utilizing Python and PySpark, demonstrating expertise in handling complex data processing and transformation workflows.
- Achieved a 20% improvement in database performance through SQL query optimization and implementation of efficient data retrieval and aggregation strategies.
- Architected and deployed comprehensive data warehouse solutions on AWS, showcasing strong proficiency in cloud-based data infrastructure.
- Engineered robust data integration workflows to facilitate seamless data movement across multiple systems and platforms, enhancing overall data accessibility.
- Established and executed comprehensive data quality frameworks, including data cleansing protocols and validation procedures, to maintain high standards of data accuracy and integrity.
- Collaborated effectively with cross-functional teams to gather requirements and deliver technical solutions for data-driven initiatives, combining strong communication skills with technical expertise.

EXPERIENCE:

Santrotech Software Technologies

DATA ENGINEER -> February 2021 to February 2024

Medilenz Innovation Private Limited

DATA ENGINEER -> March 2024 - present

PROJECTS:

Data Lake Automation with AWS Glue for Sales optimization

Develop an automated data lake solution on AWS using Python and services like S3, Glue, and Redshift, enabling seamless data ingestion, processing, and querying. The pipeline ingests data from various data sources, performs transformations, and loads the processed data into Redshift data warehouse for analysis.

Customer Transaction Data Pipeline

Develop a data pipeline to process customer transaction data from various sources. The pipeline will ingest data from AWS S3, transform it using PySpark, and store the processed data in an AWS RDS (MySQL) database. The data will be used to generate insights and reports for business analysis.