

Soni Badu

📍 sonibadu17@gmail.com 📞 +91-9337528787 🌐 <https://www.linkedin.com/in/soni-badu>

SUMMARY

Accomplished Data Engineer with over 3 years of hands-on experience, adeptly versed in a spectrum of competencies including Hadoop, Unix, API integration, data pipeline architecture, and SQL proficiency. Demonstrated ability in the execution and sustenance of Hadoop-driven solutions tailored for the management and processing of extensive data volumes.

SKILLS

Python(Pandas)

Snowflake , Oracle, MSQl

familiar with & have worked on:- Unix, Apache Spark, Kafka, Airflow ,Hive , AWS-EMR

ROLES & RESPONSIBILITIES

Technology Analyst (Data & Analytics)

Infosys Limited @Apple Client

April 2023 - Present, Bangalore

- Built scalable data pipelines using Apache's large-scale processing systems, achieving a 30% improvement in data processing efficiency for handling diverse structured and semi-structured data sources. Utilized Python (Pandas DataFrames), Spark, Kafka, and Airflow for optimal pipeline construction.
- Led data modelling efforts and contributed to Big Data technology design, resulting in a 25% reduction in data processing time and enhanced data solution robustness.
- Successfully conducted Unit, Integration, and User Acceptance Testing for data projects, achieving a 20% increase in data quality and reducing post-deployment issues by 15%.
- Expertly harnessed APIs to extract and seamlessly integrate data, enhancing pipeline efficiency by 40% and reducing data integration time by 25%.
- Conducted root-cause analyses for production challenges, leading to a 20% decrease in system downtime and faster resolution of critical issues. Played a pivotal role in migrating pipelines from Oracle to cloud-based Snowflake, resulting in a 30% reduction in data transfer latency and improved data accessibility.
- Contributed to post-production knowledge transfer, facilitating a smooth transition of pipelines, and ensuring a 95% retention of critical data architecture insights for the team.

Senior Systems Engineer (Data & Analytics)

Infosys Limited @Apple Client

April 2021 - March 2023, Hyderabad

- Developed and maintained Spark-based pipelines to facilitate smooth data transfer from source databases(Oracle , MSQl) to the S3 cloud, with subsequent loading into Snowflake.
- Showcased adeptness in refining SQL queries, resulting in amplified query efficiency and decreased execution durations.
- Significantly contributed to the knowledge transfer process during the post-production deployment phase, guaranteeing the seamless progression of pipelines.
- Increased data ingestion rate by 30% while increasing data accuracy by 10%.Optimized pipeline performance and stability with monitoring solutions as per business requirement.

Systems Engineer (Data & Analytics)

Infosys Limited @Apple Client

August 2020 - February 2022, Hyderabad

- Built the semantic pipelines post that to increase the data proficiency in Airflow using Python.
- Implemented technologies like Spark, EMR for overall efficiency, enabling more informed decision-making and improved business outcomes.
- Collaborated with the Business Analytics team for effectively delivering requirements.

CERTIFICATIONS

Infosys Certified Snowflake Data Warehouse

Infosys Certified Spark Professional

Infosys Certified Agile Developer, Data Management Associate & Business Intelligence Associate

Infosys Certified Big Data Professional

Infosys Limited • 2023

Insta Awards & DNA Rise Awards by Infosys

Infosys Limited • 2022

INVOLVEMENT

Technology Analyst (Data & Analytics)

Infosys Limited

- Served as a Team Lead for a 10-member team leading multiple data migration projects. Onboarded and conducted KT sessions for 10 new team members, resulting in a 100% accuracy in the project deliverables..
- Led the migration of 200+ interfaces from Teradata & Oracle to Snowflake platform.
- Planned and implemented Snowflake cost optimization and performance enhancement measures.
- Evolved as a key contributor and team leader through extensive stream and project training; improved team's efficiency by 40% and secured the project within budget. Automated data loading processes using Python, SQL, and Bash scripting, leading to an overall 30% improvement in data load time.

EDUCATION

Bachelor of Technology in Electrical Engineering(Honours)

Parala Maharaja Engineering College • 2019