SANTHOSH RK

SKILLS

PROGRAMMING

Python
SQL

DATA

Hadoop (HDFS, Hive)   Sqoop
ETL Pipelines   PowerBI
Spark (PySpark & SparkSQL)
PostgreSQL MySQL

CLOUD

Azure Datafactory   Azure Data Lake
Azure Databricks   Azure Synapse
Azure SQL Database Azure Keyvault
Azure Blob Storage   Azure HD Insight
AWS S3

EDUCATION

BSc in software systems   ,
Sri Krishna Arts and science college

2017-2020   Coimbatore

Certifications

Microsoft Certified Azure data engineer associate DP-203

Microsoft certified data fundaments Dp-900

Infosys certified Spark Developer

Infosys certified Big data developer

Data Engineer

Phone- 8072317325   Email-santhosh@gmail.com
Bengaluru https://www.linkedin.com/in/santhosh-rk

EXPERIENCE

Azure Data Engineer, Feb23 to present   Domain: Energy services

PROJECT DESCRIPTION: Project architecture involves data flow from Access DB to Azure sql data base using Data factory.

RESPONSIBILITIES:

*Designed and implemented ETL processes using Azure Data Factory, resulting in a 50% reduction in processing time.

*Designed a pipeline that Migrates data from **Access dB to Azure SQL** database

*Implemented **self hosted IR** for migration from ON Prem to azure cloud

*Gathering requirement from client and documented Every Operation for the migration

Azure Data Engineer, Infosys Ltd   2021 Jan to Jan 2023 Domain: Banking

PROJECT DESCRIPTION: On a daily basis, the project involves gathering the transaction details of customers from the source, cleaning and transforming them according to the business requirements, and providing them to the analytics team for reporting.

RESPONSIBILITIES:

* Designed, developed, and deployed pipelines using Azure Services **(Data Factory, Databricks, PySpark, SQL).** • Maintained and scaled data pipelines while building new data source integrations to support increasing data volume and complexity.

* **Optimized existing pipelines** resulting in a **10x** reduction in execution time and implemented cost-optimization strategies to reduce cloud costs.

* Troubleshot and resolved performance issues, optimized queries, and developed complex SQL queries to retrieve data from the database.

* Automated pipeline execution using various types of triggers and integration runtime, monitored pipelines daily, and troubleshooted issues if the pipeline failed.