

Azure Data Engineer
Name: Mukilan K L

Email: mukilankl8@gmail.com
Mobile No: +91 7339181947
[linkedin.MukilanDataEngineer](#)

Professional Summary:

- Data Engineer with 2.6 years of experience in Analysis, Development and Implementation of business applications using Azure and On-prem Environment.
- Strong Experience in Azure Data Factory, Azure Databricks, Azure Key Vault, Azure Logic App, Azure Data Lake, Azure Storage, Azure SQL.
- Experience in extraction of data from various Heterogeneous sources (Relational database, SAP ERP, API's and sFTP server) into Data Warehouse/Datamart targets.
- Having good knowledge on Hadoop and Spark Architecture and Hands-On experience in file format (parquet, orc, avro), CI/CD process, Optimization and compression technique.
- Proficient in understanding business processes/requirements and translating them into technical requirements.
- Excellent interpersonal and communication skills, customer skills, documentation skills and desire to work in dynamic and challenging environments.
- Work exceptionally well in a team or in an individual environment (self-starter) and ability to learn new technologies and tools quickly.

Professional Experience:

- Currently working with TheMathCompany as a Azure Data Engineer from DEC 2022 to till now, Bangalore.
- Worked in Enquero as a Azure Data Engineer from JULY 2022 to NOV 2022, Bangalore.
- Worked in Cognizant Technology and Solutions as a Programmer Analyst from FEB 2021 to JULY 2022, Chennai.

Educational Qualification:

- B. Tech from Sri Manakula Vinayager Engineering College 2020, Pondicherry.

Technical Skills:

Databases	Azure SQL, DBeaver
Big Data Ecosystem	Spark, Hive, Hadoop, Cloudera, Sqoop
Programming Languages	Python, Pyspark
Cloud Skills	Azure Data Factory, Azure Databricks, Azure Storage account, Azure Logic App, Azure Key Vault.
Platform Skills	JIRA, Azure Board, Azure DevOps, Hadoop Linux

Certification and Project Achievement:

- AZ - 900 Azure Fundamentals
- DP - 900 Azure Data Fundamentals
- Databricks Certified - Lakehouse Fundamental
- AWS Essentials – Udemy
- Received many appreciations from clients.

Projects Details:

Project#3:

Client : Bridgestone
Designation : Azure Data Engineer
Team Title : Demand Forecast DnA
Duration : Dec 2022 till now
Domain : Manufacturing
Environment : Azure Data Factory, Azure Databricks, Azure Key vault, Azure DevOps

Roles and Responsibilities:

- Collaborate with cross-functional team to identify data requirement and develop data models.
- Built and maintained efficient azure data factory pipeline to extract, transform, and load data from multiple sources into azure blob storage and azure data lake storage Gen2.
- Expertise in creating generic pipeline for data ingestion and transformation.
- Implemented metadata framework for data pipeline and followed best practice.
- Used Pyspark in databricks for data extraction, transformation, and aggregation from multiple file formats.
- Transformed and cleaned the dataset and hive tables in databricks notebooks with SQL, Pyspark to generate the core layer in data warehouse tables and storing them in delta format and using merge to upsert the data into the data lake tables.
- Implemented mail alert mechanism using Azure logic app.
- Proficient in working with databricks delta format.
- Involved in creating medallion architecture by leveraging databricks lakehouse platform.
- Understanding the functional aspects of the clients requirement and efficient task execution.

Project#2:

Client : GE Power
Designation : Azure Data Engineer
Team Title : DnA - GE
Duration : JULY 2022 to NOV 2022
Domain : Supply chain and logistics
Environment : Azure Data Factory, Azure Databricks, Azure Key vault, Azure DevOps

Roles and Responsibilities:

- Designed and developed data pipeline by using Azure data factory for Data Migration that extract data from the heterogeneous systems, and transform it as per business requirements.
- Error Handling Mechanism and Store the Log Information.
- Implemented optimization technique and scheduling notebooks with data factory by applying dependencies between notebooks and passing the parameters.
- Data validation was performed to ensure that all value are valid.
- Used Data Frames for the analysis by taking the source data from files and tables, using SparkSQL for transformations then loading the target table or files based on business requirements.
- Make sure that all dataset are loaded into the data warehouse without any data loss and truncation
- Developed scalable data pipeline using Fast/Slow lane data.

Project#1:

Client : Williams-sonoma
Designation : Data Engineer
Team Title : WSGC - Support
Duration : MAR 2021 – JULY 2022
Domain : Supply chain and logistics
Environment : Hive, Sqoop, Pyspark, Cloudera, Impala

Roles and Responsibilities:

- Worked in scrum team environment with scrum master, product owner and peers to develop and deploy logic.
- Load and Transform large data sets of structure from Relational Database System to HDFS and vice-versa using Sqoop tool.
- Designing and creating Hive external tables and creating Partitions and Bucketing techniques in Hive using Hive Query Language (HQL).
- Implemented Partitioning and Bucketing on Hive table for better performance.
- Developed and enhanced spark and hive scripts with better optimization
- Make sure that the count of records loaded in the target is matching with the expected count.
- Verify that data is transformed correctly according to various business requirements and rules.
- Responsible for enhancing the existing code with respect to stakeholder's requirements.

Strengths and Personal traits:

- Quick learner.
- Diligent and smart working.
- Ability to resolve challenging technical issues.
- Flexible with any type of technology and work environment.
- A high tolerance of stress and enjoys responsibilities
- Good team player with positive attitudes.
- Good in email communication with perfect email etiquette.