

V Mahesh

Data Engineer

07818953549 | vmahi.026@gmail.com | www.vegumaheshbabu.com | linkedin.com/in/vegumaheshbabu | London

SUMMARY

Data Engineer with expertise in transforming complex datasets into actionable insights. Skilled in leveraging Azure tools like Databricks, Data Factory, and SQL for data collection, processing, and analysis. Proficient in PySpark, SQL, Apache Spark, Apache Kafka, exploratory data analysis, and statistical data analysis. Proven track record of leading data engineering projects, collaborating across teams, and delivering high-impact solutions. Eager to embrace new challenges in a fast-paced, growth-focused environment

SKILLS

Programming Languages: Python, SQL, Pyspark, Spark

Cloud Services and Technologies: Azure Databricks, Azure Data Factory, Azure Synapse, Apache Kafka, Data modelling, Azure Synapse, Data Normalization and warehousing, CI/CD Pipelines

EXPERIENCE

GMK Software Sols, London UK

June 2023 – Sep 2024

Data Engineer

- Designed and developed **ETL** architecture and **Data Pipelines** in **Azure Cloud** from scratch with **Java**, **Python**, and **SQL**, enabling efficient loading of operational data into the **Data Warehouse** and achieving a **99.9%** accuracy rate for loaded data.
- Designed and implemented a comprehensive data cleansing framework using **Python**, and **SQL**, resulting in a **60%** reduction in manual effort required for data cleaning.
- Built Azure Databricks notebooks and built business models using SQL, Python, and Java to provide business owners with greater visibility on terminal performance, enhancing decision-making processes.
- Collaborated exclusively with **stakeholders** from 5 departments to meticulously gather requirements, facilitating the design, construction, and execution of a spectrum of data platform initiatives.
- Pulled data from multiple sources, including **Rest API**, **Kafka**, **SQL Server**, **Blob**, **Gen2**, and **Delta Lake**, using **ADF**, **Databricks**, and **Java**.

Bank Of America

Jan 2022 – Dec 2022

Data Engineer

- Worked exclusively with stakeholders across 6 departments to gather requirements, design, build, and execute various data platform initiatives.
- Engineered multiple data pipelines using **Databricks** and **Azure Data Factory (ADF)**, extracting data from over **10** structured and unstructured sources. Additionally, created **triggers** to automate **Data Factory pipelines**, **Dataflows**.
- Orchestrated seamless data ingestion and storage across multiple sources into **Azure Data Lake Storage (ADLS) Gen2**, amplifying data accessibility by **30%**.
- Extracted business and finance data from various sources, including **Azure SQL**, **Snowflake**, **Amazon Redshift**, on-premises SQL Servers, and various web applications, and staged it in ADLS Gen2 using Azure Data Factory (ADF) pipelines.
- Assisted more than **100** end users in troubleshooting data-related issues, performance issues resolving queries, and ensuring the smooth functioning of BI Sols.
- Developed Power BI reports to analyze business trends, improving decision-making efficiency by **20%** and increasing quarterly sales by **15%**.

Accenture

June 2021 – Jan 2023

Data Engineer

- Collaborated with **stakeholders** from various departments, including business users, IT teams, and executives to understand data requirements and translating them into **10+** business dashboards.
- Create and implement an ETL pipeline using **Azure Databricks** and **Data Factory** to automatically extract data from various sources and prepare it for reporting teams by cleaning data, de normalizing, and organising it.
- Led the successful migration of complex company data from on-premise to cloud using **Databricks** and **Azure Synapse**, Leading to a **25%** increase in performance.
- Established frameworks for **data governance** to guarantee the appropriate management, security, and privacy of data.
- Updating current dashboards, reports, and **dataflows** while focusing on ad hoc tasks to enhance such reports.
- Technically supported internal teams in troubleshooting and resolve **300+** data-related difficulties.

Claro software solutions

July 2020 – May 2021

Data Analyst

- Worked and interpreted diverse client finance and sales datasets to identify trends and insights.
- Delivered bi-monthly reports to upper management, providing strategic insights for decision making, and presented clear data insights to business stakeholders.
- Using Power BI, generated **15+** story-telling dashboards that deliver actionable insights into operational efficiency.
- Published reports to Power BI Service, managed users and user groups, implemented Row-Level Security (RLS), and configured scheduled data refreshes using Gateways for on-premises data sources

EDUCATION

Master's in data science

University of Hertfordshire

Jan 2023 – June 2024

Hatfield, United Kingdom

Bachelor of technology in Computer Science

Gitam university

June 2017 – June 2021

Visakhapatnam, India

CERTIFICATIONS

DP-900 Azure Data Fundamentals

DP 203 – Azure Data Engineer Associate

COF-CO2- Snow Pro core Certification

AWS Cloud Practitioner

Python Specialization from University of Michigan