

GURUDEV H Y

DATA ENGINEER

✉ gdev701@gmail.com 📞 +91-9972039190 📍 Bengaluru , Karnataka, IN
in <https://www.linkedin.com/in/gurudev-h-y-947913174/> 🌐 <https://medium.com/@gdev701>

SUMMARY

Experienced Data Engineer with nearly 2 years of hands-on expertise in designing, implementing, and maintaining scalable data pipelines and analytical solutions. Proficient in utilizing various technologies and programming languages to extract, transform and load (ETL) data from diverse sources into data warehouses and lakes. Skilled in optimizing data infrastructure for performance, reliability, and efficiency, ensuring seamless data flow and accessibility for stakeholders.

SKILLS

Python	● ● ● ● ○	Azure DataLake	● ● ● ○ ○
SQL Queries	● ● ● ● ○	Azure SQL Database	● ● ● ○ ○
Microsoft SQL	● ● ● ● ○	Data-warehousing and Data modeling	● ● ● ○ ○
Azure Data Factory	● ● ● ● ○	Apache Airflow	● ● ● ○ ○
Azure Databricks	● ● ● ● ○	Azure Devops	● ● ● ○ ○
Azure Synapse Analytics	● ● ● ○ ○	Azure Logic Apps	● ● ○ ○ ○

WORK EXPERIENCE

Data Engineer

TVS MOTOR COMPANY LTD *Jul 2021 - Jan 2023*

- Developed and deployed ETL/ELT processes to streamline data ingestion, transformation, and loading tasks, resulting in improved data quality and accessibility.
- Utilized tools such as Apache Spark, Apache Airflow, and Azure Data-factory to orchestrate and automate data workflows, enhancing operational efficiency.
- Utilized Azure Data bricks to develop and orchestrate Extract-Transform-Load (ETL) processes for ingesting data from on-premises databases, cloud-based applications, and APIs.
- Leveraged Apache Spark for distributed data processing, enabling parallel execution of data transformations and optimizations for performance.
- Designed and implemented a modern data warehouse solution using Azure Data bricks to enable interactive querying and analytics.
- Collaborated with cross-functional teams including analysts, and software engineers to understand data requirements and deliver scalable solutions that meet business objectives.
- Continuously learning and adapting to emerging technologies and best practices in the field of data engineering to drive innovation and efficiency.

PROJECTS

Vehicle Manufacturing Data Dashboard

Data Engineer

- Objective:** Developed a comprehensive data pipeline to monitor and analyze vehicle manufacturing data, enabling real-time insights into production performance and operational efficiency.
- Tools and Technologies:** Apache Airflow, MongoDB, Apache Spark, SAP HANA, Python
- Key Responsibilities:**
 - Designed and implemented ETL processes using Apache Airflow to orchestrate data ingestion from MongoDB, ensuring timely and accurate data flow.
 - Processed and transformed manufacturing data using Apache Spark to handle large volumes of data efficiently.
 - Created data models in SAP HANA to organize and store processed data, facilitating efficient querying and reporting.
 - Conducted data validation and quality checks to ensure the accuracy and reliability of the data.

- Collaborated with cross-functional teams to gather requirements and ensure the data pipeline met business needs.
- **Senior Contribution:**
 - Developed interactive and dynamic dashboards using MicroStrategy to visualize key performance indicators (KPIs) such as production volume, defect rates, and downtime, based on the data provided by the pipeline.
 - Automated data refresh and reporting schedules to provide up-to-date insights to stakeholders.

Vehicle Quality Dashboard

Data Engineer

- **Objective:** Developed a data pipeline and dashboard to monitor and analyze quality metrics within the manufacturing process, enabling real-time insights and improved quality control.
- **Tools and Technologies:** Microsoft SQL Server, Azure Data Factory, Azure Data Lake, Azure Databricks, Python, Apache Spark
- **Key Responsibilities:**
 - Designed and implemented ETL processes using Azure Data Factory to orchestrate data ingestion from Microsoft SQL Server into Azure Data Lake, ensuring timely and accurate data flow.
 - Processed and transformed quality data using Azure Databricks with pyspark, handling large volumes of data efficiently.
 - Created data models in Azure Databricks to organize and store processed data, facilitating efficient querying and data warehousing.
 - Collaborated with cross-functional teams to gather requirements and tailor dashboard features to meet business needs.
 - Conducted data validation and quality checks to ensure the accuracy and reliability of the dashboard data.
- **Achievements:**
 - Employed Azure Databricks for data processing, optimizing Spark jobs to enhance performance.
 - Applied Spark optimizations such as partitioning and caching to reduce processing time by 50%.

EDUCATION

Bachelor of Technology in Computer Science and Engineering

Dayannada Sagar University *Aug 2017 - Jun 2021*

CERTIFICATIONS

Microsoft Certified: Azure Data Engineer Associate (21st May 2024 - 21st May 2025)

Skills measured:

- Design and implement data storage
- Develop data processing
- Secure, monitor, and optimize data storage and data processing

BREAK

Career Break (13th Jan 2023-current)

- Planned to pursue higher studies but had to drop it, due to caregiving responsibilities for a family member's health issues. During this period, I focused on providing familial support and independently prepared for and successfully completed the Microsoft Certified: Azure Data Engineer Associate (DP-203) certification. This certification enhanced my expertise in Azure cloud technologies and data engineering principles, preparing me for a career transition back into the workforce.