

# YASHASVI SHARMA

Azure Data Engineer | ETL Developer | Databricks Engineer

 +1 (289)-236-6906

@ [sharmayashasvi876@gmail.com](mailto:sharmayashasvi876@gmail.com)

 [LinkedIn](#)

 Toronto, Canada

 [GitHub](#)

## EDUCATION:

Bachelors of Science (Honours),  
Computer Science with Minors in Business,  
Brock University, St. Catharines, Canada.

Sept 2019 – Dec 2023

## CERTIFICATIONS:

- DP-700: Microsoft Fabric Data Engineer Associate
- DP-900: Microsoft Certified: Azure Data Fundamentals
- Google Data analytics

## PROFESSIONAL SUMMARY:

- Highly motivated Azure Data Engineer with 1+ year of experience designing, developing, and deploying **data pipelines** and data warehousing solutions on the **Microsoft Azure platform**
- Proven ability to leverage **Azure services** like **Azure Data Factory**, **Azure Databricks**, **Azure Synapse**, and Storage solutions to build scalable and efficient data architectures
- Possess in-depth knowledge of **Azure Data Factory (ADF)** for orchestrating complex data workflows, **Azure Databricks** for advanced data processing with **PySpark**, and **Azure Synapse Analytics** for unified data warehousing and analytics
- Skilled in designing and implementing data models for efficient data storage and retrieval, Utilize **ETL/ELT** processes to seamlessly move data between diverse sources and destinations
- Leverage **Azure Data Lake Storage (ADLS)** for scalable **data lakes**, **Azure Blob Storage** for unstructured data management, and **Azure SQL Database** for relational data storage, ensuring secure and reliable data access
- Adept at utilizing **Azure Stream Analytics** to process and analyze real-time data streams to extract valuable insights from continuous data streams, enhancing operational efficiency and responsiveness
- Experience in implementing data governance and security measures using **Azure Key Vault**, **Unity Catalog**, and **Role-Based Access Control (RBAC)** to protect sensitive data
- Implemented Proactive data monitoring practices using **Azure Monitor** to ensure data pipeline health and identify potential issues, minimizing downtime and optimizing data processing efficiency
- Possess strong programming skills in **Python** and **PySpark** to manipulate and analyze data efficiently, enabling streamlined data processing workflows and advanced analytics capabilities
- Basic Understanding ability to troubleshoot performance bottlenecks and optimize query execution using **SQL** tuning techniques in **Snowflake** and **Azure Synapse Analytics**
- Successfully migrated data from various sources to target platforms, ensuring data integrity and continuity, and consistency and deliver migration projects within scope, budget, and timeline constraints
- Good in integrating **Jira** with **Confluence** and collaborating with cross-functional teams for Agile project management documentation and issue tracking

## SKILLS:

- |                           |                          |                     |
|---------------------------|--------------------------|---------------------|
| • Azure Blob storage      | • Azure SQL Database     | • Python            |
| • Azure Cosmos DB         | • Azure Stream Analytics | • PySpark           |
| • Azure Data Bricks       | • Azure Synapse          | • SQL               |
| • Azure Data Lake Storage | • Data Pipelines         | • Ms Fabric         |
| • Azure Data Factory      | • Snowflake              | • Power BI          |
| • Azure Key Vault         | • Azure Monitor          | • Jira & Confluence |

**TECHNICAL SUMMARY:****Azure Data Engineer,****Nestle, Canada. Mar 2024 – Till date**

- Designed and implemented data ingestion pipelines using Azure Data Factory to extract, transform, and load (ETL) data from various sources into Azure Blob storage and Azure Data Lake Storage
- Developed and maintained data processing workflows using Azure Databricks and PySpark to handle large-scale data transformations and analytics tasks efficiently
- Created and managed data catalogs using Azure Data Catalog to enable easy discovery and collaboration on data assets across the organization
- Implemented data security and access controls for Azure SQL Database and Azure Cosmos DB to ensure compliance with regulatory requirements
- Conducted performance monitoring and tuning of Snowflake data warehouses, including workload management, query optimization, and resource scaling, to achieve optimal performance and cost-effectiveness
- Orchestrated end-to-end ETL pipelines using PySpark and Snowflake to extract data from various sources, transform it according to business logic, and load it into data warehouse
- Utilized Azure Synapse Analytics for building and optimizing data warehousing solutions, including data modeling, partitioning, and indexing for improved query performance
- Collaborated with stakeholders to gather requirements and design scalable data architectures for supporting business intelligence and reporting needs in Power BI
- Automated data integration processes and workflows using Python scripts and Azure Data Factory pipelines to improve operational efficiency and reduce manual effort
- Monitored and optimized data pipelines and storage solutions using Azure Monitor, ensuring high availability, performance, and reliability of data services