

Bhavya Appichi | Azure Data Engineer | Databricks Data Engineer

✉ abhavya155@gmail.com | ☎ +91-9121397752 | 🔗 [LinkedIn](#) | 💻 [GitHub](#) | 🌐 [Portfolio](#)

Professional Summary

Data Engineer with **3.5 years of hands-on experience**, including **2.5 years of professional work** at Cognizant and **1 year of upskilling** through self-driven, project-based learning. Proficient in designing and implementing data pipelines and data integration solutions by using SQL, Pyspark. Processing and analyzing over 1 million records daily, resulting in a 25% improvement in analysis accuracy. Worked with cross-functional teams to deliver high-quality data solutions. Maintains a strong GitHub portfolio and a live technical website to showcase end-to-end projects.

Technical Skills

- **Cloud Platforms:** Azure Data Factory, Azure Databricks, Azure Data Lake Storage Gen2, Azure SQL Database, Azure Blob Storage, Azure Event Hubs, Azure Key Vaults, Azure Logic App
- **Programming Languages:** SQL, PySpark
- **Big Data Technologies:** Apache Spark, Delta Lake
- **Data Management:** Data Warehousing, ETL Pipelines, Data Lakes, Data Governance & Security
- **Tools & Platforms:** Azure DevOps, GitHub, Azure Repos
- **Databases:** SQL Server, Oracle

Career Gap

Professional Upskilling | Apr 2024 - Present

Focused on deepening data engineering skills. Completed 3 end-to-end Azure projects, contributed to open-source datasets, and published a technical portfolio showcasing capabilities in real-time, batch, and hybrid data pipelines.

Professional Experience

Data Engineer

Cognizant Technology Solutions | Nov 2021 - Mar 2024

Associate - Travelers impaired Hours | Aug 2023 – Mar 2024

- Developed and optimized scalable ETL pipelines using Azure Data Factory and Azure Databricks to process 1TB+ datasets, resulting in a 30% increase in data reliability and consistency
- Designed and implemented data pipelines using the medallion architecture (Bronze to Silver to Gold) in Databricks, improving data reliability by 30%.
- Integrated Databricks with Azure Data Lake Storage and enterprise data warehouses to enable centralized data management and improve accessibility and processing efficiency by 40%.
- Optimized data processing performance by 70% through Z-Ordering, partitioning, and file compaction techniques, enhancing overall efficiency.
- Delivered incremental batch processing in Databricks using Delta Lake, improving data accuracy and pipeline efficiency by 25%
- Processed scalable data transformation workflows using Apache Spark, Python, and SQL, implementing a Lakehouse architecture that reduced report generation time by 25%.
- Applied data governance and compliance standards, ensuring 99% data quality.
- Leveraged Azure DevOps to create CI/CD pipelines, cutting deployment time by 50% and enhancing the reliability of data pipeline releases.

Programmer Analyst - PetSmart Data Integration | Nov 2022 - July 2023

- Developed a **Metadata-Driven Incremental Loading Pipeline** using Azure Data Factory (ADF), optimizing data processing workflows and achieving a reduction in processing time by 50%.
- Configured secure **ADLS mounting** in Azure Databricks using Azure Key Vault **Service Principle**, ensuring 100% role-based access control (**RBAC**) and data security
- Enhanced data logging for ADF pipelines, improving execution tracking accuracy by 70% leading to more informed decision-making across teams.
- Engineered automated email notifications for pipeline failures, enhancing response time by 50% through integration with Azure Logic Apps.
- Automated data ingestion jobs in Azure Data Factory, reducing manual effort by 80% through the implementation of scheduled triggers

Programmer Analyst Trainee - Travelers Audit Data Integration | Nov 2021 - Oct 2022

- Pioneered batch and real-time pipelines to process 1M+ daily records into Azure Data Lake, boosting data availability and reporting speed by 30%.
- Engineered 10+ parameterized pipelines to ingest and process over 500,000 records daily from REST APIs into ADLS Gen2, reducing data ingestion time by 35% and improving pipeline reliability.

Highlighted Projects

- **Azure Data Migration Project**
Migrated data from on-premises SQL Server, Azure Blob Storage, and REST APIs to Azure SQL Database and Azure Data Lake Storage (ADLS) using Azure Data Factory (ADF). [GitHub](#)
- **Databricks Incremental Data Pipeline Project**
Orchestrated a production-ready Databricks pipeline leveraging Delta Lake (Bronze-Silver-Gold). the pipeline featured data validation and comprehensive logging for auditing. [GitHub](#)
- **Weather Heatwave Alert Streaming Project**
Built with Azure Event Hub, Databricks, and Delta Lake. Streamed 10K+ events/hour with <5s latency. Used SCD Type 2 for dimension tracking. [GitHub](#)

Education

Bachelor of Technology in Electrical and Electronics Engineering
JNTUA College of Engineering, Pulivendula | Jul 2017 - Jun 2021

Certifications

- AWS Certified Cloud Practitioner

Achievements

- Delivered scalable cloud-native solutions with 25–50% improvements in efficiency.
- Built 3 high-quality projects with public GitHub code and live portfolio.