# Teja C

## Data Engineer

London, UK                                      +44 7776632365

[teja9509@gmail.com](mailto:teja9509@gmail.com)                    [www.linkedin.com/in/vtc99/](http://www.linkedin.com/in/vtc99/)

## Professional Summary

Certified Data Engineer with 4+ years of experience in building and optimizing data platforms using Microsoft Azure. Specialized in ETL pipelines, lakehouse architectures, and real-time analytics solutions using **ADF, Databricks, Synapse, and Delta Lake.** Hands-on experience on **AWS(EC2, Glue, S3).** Adept at handling large-scale structured and semi-structured data with Apache Spark and PySpark, implementing **CI/CD** pipelines, and ensuring data quality, lineage, and governance across enterprise environments.

## Technical Skills

- **Cloud & Data Platforms**: Azure Data Factory, Azure Synapse, Azure Databricks, Azure Data Lake Gen2, Azure SQL, AWS Glue, AWS S3, AWS EC2.

- **Data Engineering & Processing**: PySpark, Delta Lake, Apache Spark, Structured Streaming, Azure Event Hubs.

- **ETL & Workflow Orchestration**: ADF, Mapping Data Flows, Notebooks, Azure Functions, Logic Apps.

- **Data Modeling & Analytics**: Star/Snowflake Schema, Power BI.

- **Security & Governance**: RBAC, Azure Key Vault, Data Lineage, Sensitivity Labels, Auditing.

- **DevOps & CI/CD**: Git, Azure DevOps, Pipelines, Log Analytics, Monitoring & Alerts.

- **File Formats**: CSV, JSON, Apache Parquet.

- **Languages :** Python, PySpark, SQL

- **NoSQL:** MongoDB, Cassandra, Neo4J.

## Professional Experience

### Azure Data Engineer
Chubb Ltd (2023 – Present)

- Designed and implemented scalable ETL pipelines using Azure Data Factory and Databricks to process multi-format data (CSV, Parquet, JSON) into Delta Lake.
- Architected and deployed lakehouse infrastructure using Azure Data Lake Gen2, Delta Lake, and Synapse Analytics for high-performance query workloads.
- Integrated Azure Functions and Logic Apps for event-driven orchestration and real-time processing.
- Created Power BI datasets and semantic models using Synapse SQL for enterprise reporting.
- Developed CI/CD pipelines with Azure DevOps for automated deployment of data services and notebooks.
- Set up monitoring and alerting with Azure Monitor, Log Analytics, and diagnostic settings.
- Ensured compliance with security standards using RBAC, Key Vault, and data masking.

### Data Engineer – Cloud & Big Data
Wipro Ltd (2020 – 2022)

- Migrated legacy SSIS workflows to modern ADF pipelines, using Mapping Data Flows and Azure IRs.
- Processed large-scale transactional datasets (2TB+/day) using Databricks and PySpark on Azure Data Lake.
- Designed dedicated SQL pools in Synapse Analytics for optimized querying and reporting.
- Created modular notebooks for transformation logic and schema evolution with Delta Lake.
- Developed Power BI dashboards from Synapse datasets for operational and business users.
- Integrated Azure Key Vault and managed identities for secure pipeline authentication.
- Implemented data governance practices: lineage tracking, auditing, and access control.

## Key Projects

### Enterprise Data Lakehouse (Chubb)
- Built a Delta Lake–based lakehouse architecture on Azure Data Lake Gen2 with ADF + Databricks, supporting both batch and streaming data ingestion.
- Enabled ACID-compliant data operations and real-time data availability for BI dashboards.

### Real-Time Processing & Alerting (Chubb)
- Developed event-driven pipelines using Event Hubs, Azure Functions, and Data Factory, enabling near real-time alerts for incoming data anomalies.
- Integrated with Azure Monitor and Log Analytics for proactive system health reporting.

### ETL Modernization & Performance Tuning (Wipro)
- Migrated SSIS workflows to cloud-native ADF pipelines, improving transformation speed by 60%.
- Reduced cost and increased reliability by leveraging serverless Spark pools and partitioned views.

### Enterprise Data Warehouse (Wipro)
- Designed and built Synapse-based warehouse with star schema, integrated with Power BI for dynamic, role-based dashboards across departments.

## Education

MSc. Data Science

University of Greenwich, London – 2022 – 2023

Relevant Modules: Big Data Engineering, Cloud Platforms, Real-Time Data Processing, SQL, MongoDB, Cassandra, Neo4J, Spark & PySpark.

## Certifications
- Microsoft Certified: Azure Data Engineer Associate (DP-203)
- Designing & Implementing Data Solutions with Azure Databricks –Udemy