Saurabh Dubey

Junior Data Engineer

☎ +919503371603 | ✉ dubeysaurabhrajendra@gmail.com | ⌨ SaurabhRajendra7584 | in saurabhrajendradubey

PROFESSIONAL SUMMARY

Entry-level Azure Data Engineer with hands-on experience in building and supporting cloud-based data pipelines using Microsoft Azure. Strong foundation in ETL/ELT processes, data ingestion, transformation, and analytics-ready data modeling. Practical exposure to Azure Data Factory, Azure Data Lake Storage, Azure Databricks, and Azure Synapse Analytics through personal projects. Looking to contribute to compay's data engineering initiatives while continuously learning enterprise-scale data platform best practices.

TECHNICAL SKILLS

Cloud Platform: Microsoft Azure

Azure Data Services: Azure Data Factory, Azure Data Lake Storage (ADLS Gen2), Azure Databricks, Azure Synapse Analytics, Azure SQL Database, Azure Blob Storage

Programming & Query Languages: SQL, Python,PySpark

Data Engineering Concepts: ETL / ELT Pipelines, Batch Data Processing, Data Ingestion & Transformation, Data Modeling, Incremental Data Loads, Data Validation & Quality Checks

Big Data & Processing: Apache Spark (Foundational)

Dev & Collaboration Tools: Git, GitHub, Azure DevOps (basic CI/CD exposure), Agile / Scrum

PROFESSIONAL EXPERIENCE

Business Analyst Intern   GRT Global Logistics Pvt. Ltd.   Nov 2025 – Jan 2026

- Gathered, documented, and translated business requirements into functional specifications and user stories
- Collaborated with stakeholders to define scope, priorities, and acceptance criteria
- Managed tasks, sprint planning, and tracking using JIRA under Agile methodology
- Coordinated between business and technical teams to track dependencies, mitigate risks, and resolve issues
- Supported timely project delivery through requirement validation and progress monitoring

PROJECTS

Azure Batch Data Pipeline for E-Commerce Analytics

- Designed and implemented an end-to-end batch ETL pipeline on Microsoft Azure
- Ingested data from Azure Blob Storage and REST APIs using Azure Data Factory
- Performed data cleansing and transformation using Azure Databricks (PySpark)
- Stored processed data in Azure Data Lake Storage Gen2 following raw, processed, and curated layers
- Loaded curated data into Azure Synapse Analytics for analytical querying
- Implemented basic monitoring and logging for pipeline execution

Technologies: Azure Data Factory, ADLS Gen2, Azure Databricks, Azure Synapse Analytics, PySpark, SQL

Streaming Data Ingestion – Conceptual Implementation

- Built a foundational real-time ingestion pipeline using Azure Event Hubs to process streaming data
- Applied transformation logic and stored processed data in Azure SQL Database for querying
- Implemented basic error handling and data validation rules
- Created simple dashboards in Power BI to visualize incoming data trends

Technologies: Azure Event Hubs, Azure SQL Database, Power BI

Data Warehouse Design using Azure Synapse

- Designed a star schema data model for analytical reporting
- Used Azure Data Factory to load and transform source data into Synapse tables
- Developed SQL queries to support reporting and analysis use cases
- Focused on data consistency, query performance, and structured storage

Technologies: Azure Synapse Analytics, Azure Data Factory, SQL

EDUCATION

Bachelor of Science – Computer Science

Thakur College of Science and Commerce, University of Mumbai | 2022 - 2025

Relevant Coursework: Database Management Systems, Data Science, Machine Learning , Fundamentals of NoSQL Database – MongoDB

CERTIFICATIONS

- IBM Data Fundamentals – IBM (2026)
- Data Analytics Job Simulation - Deloitte (2025)
- Data visualization - TCS (2025)
- Linux Administration - IIT Bombay (2025)

ADDITIONAL INFORMATION

- Strong interest in cloud data platforms and analytics engineering
- Actively improving skills in Azure-based data engineering & Machine Learning