



Malay Kumar

Data Engineer

I am a skilled data engineer with nearly 4 years of professional experience, with passion for transforming raw data into actionable insights. With expertise in designing, building and maintaining data pipelines to help the data analyst for building dashboards, I look forward to establish myself by adapting new technologies as well.

✉ malaykumar1939@gmail.com

📍 west Bengal, Purulia, India

☎ +91-8670860865

🌐 [linkedin.com/in/malay-kumar-1299631b7](https://www.linkedin.com/in/malay-kumar-1299631b7)

WORK EXPERIENCE

Data Eng Mgmt and Governance Senior Analyst

Accenture

07/2024 - Present

Achievements/Tasks

- Designed and implemented data pipelines to load historical and incremental data into staging and main tables using Databricks.
- Retrieve various (csv, parquet, txt) file format data from ADLS, Handling the Null values, perform various Business and validation Rules and load the data in landing delta table
- Developed complex views and stored Procedure in Azure SQL server to support PBI Report
- Designed and implemented Databricks Notebook using pyspark and spark sql to handle complex data manipulations, aggregations, and transformations while ensuring performance optimization.
- Designed and developed Azure Data Factory (ADF) pipelines for orchestrating end-to-end workflows, including integration with Databricks notebooks and SQL Server stored procedures.

Associate

Cognizant Technology Solutions

07/2021 - 07/2024

ETL Migration from On premise to Azure Cloud

Achievements/Tasks

- Migrate existing Jobs(5000+) from DataStage/Tivoli to Data Bricks and convert DataStage jobs into Pyspark Notebooks.
- Build Delta Lake as another layer while migrating ETL Jobs and maintain same data between Delta Lake and Teradata Vantage
- Accessing, extracting and transforming data from a variety of all sizes from ADLS Storage.
- Development of Notebook, Debugging the syntactical error of the notebook, Defect fixing(data type issue, count mismatch issue, data load issue)
- Optimized queries for performance by applying techniques like Broadcast Join, Repartitioning, and Coalescing, improving data processing speed and resource efficiency.

SKILLS

Pyspark

Azure databricks

Python

Azure Data Factory

Airflow

Azure Data Lake Storage

SQL

Azure Delta Lake

Spark SQL

ETL

Spark

Azure SQL Database

SQL Server

Data Warehousing

ACHIEVEMENTS

Cognizant Coral Award (10/2022 - 07/2024)

Awarded as the emerging talent of the year, for delivery of complex deliverables before the SLA

Accenture Spotlight Award (02/2025)

Create a significant impact for the Client

CERTIFICATES

Microsoft certified : Azure Data Engineer Associate (dp-203)

Azure Data Fundamentals (dp-900)

Dataricks Lakehouse Fundamentals

Apache Airflow Fundamentals(Astronomer)

EDUCATION

Kalinga Institute of Industrial Technology
Bhubaneswar (07/2019 - 05/2021)

Master of computer Applications

CGPA 9.17

LANGUAGES

Englisg

Full Professional Proficiency

Hindi

Full Professional Proficiency

Bengali

Full Professional Proficiency