

# Tahir Malik (Data Analyst / Engineer)

New Delhi, IN | [tahir.malik296@gmail.com](mailto:tahir.malik296@gmail.com) | +91 9971587221 | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

## EDUCATION

<b>Amity University (Gurugram, India)</b> <b>B. Tech in Information Technology (C.G.P.A – 8.45/10)</b> Relevant Coursework: Python, SQL, Big Data & Fundamentals, Economics (Minor)	<b>Jul 2017 – Apr 2021</b>
---	----------------------------

## PROFESSIONAL EXPERIENCE

<b>Data Engineer Intern, Mactores, Mumbai</b>	<b>Apr 2025 - Present</b>
<ul style="list-style-type: none"><li>Gained hands-on experience in data engineering workflows on the Databricks Lakehouse Platform.</li><li>Monitored and debugged ETL pipelines with Apache Spark &amp; Delta Lake to ensure reliability and data quality.</li><li>Learned how to implement real-time streaming pipelines using Kafka, Delta Lake, and Databricks Auto Loader.</li><li>Earned AWS Cloud Practitioner, Databricks Certified Data Engineer Associate &amp; Professional credentials during this role.</li></ul>	

## PROJECTS

### Fintech Data Migration Pipeline

- Developed scalable pipelines for migrating and transforming data of a fintech company from their traditional SQL database to Azure Data Lake, ensuring data consistency and minimal downtime.
- Implemented a medallion architecture within ADLS Gen2 using PySpark & Delta Lake to effectively manage and process data, improving data accessibility and reliability.

### E-Commerce Data Streaming Analytics

- Developed a robust data pipeline leveraging Confluent Kafka to collect e-commerce clickstream and user activity data, enabling real-time analytics capabilities.
- Combined Azure Data Lake Storage Gen2 (ADLS Gen2) with Databricks to efficiently store, cleanse, and convert raw data into formats suitable for analysis, facilitating actionable insights.

## TECHNICAL SKILLS

**Languages:** Python, SQL, C, C++

**Big data & Fundamentals:** Hadoop, Hive, Kafka, Apache Spark, Databricks, Airflow, MongoDB, Docker, Git

**Cloud Services:** Azure Data Factory, Azure Synapse Analytics, Azure Function App, Azure Data Lake Storage, Azure Cosmos DB, Azure SQL Database, GCP Dataproc, GCP Pub/Sub, GCP Cloud Storage, Snowflake, Dimension & Fact Table, BigQuery

**Soft Skills:** Communication, Analytical Thinking, Problem Solving, Team Player, Data-Driven Decision Making

## CERTIFICATIONS

- [Databricks Certified Data Engineer Professional](#)
- [AWS Certified Cloud Practitioner](#)
- [Introduction to Statistics](#)
- [Databricks Certified Data Engineer Associate](#)
- [Data Engineering with Azure](#)
- [Google Data Analytics Specialization](#)