

# Kunal Deshmukh

+91 9075155483 | [kunaldeshmukh5245@gmail.com](mailto:kunaldeshmukh5245@gmail.com) | [linkedin.com/in/kunal-deshmukh-a94b69179](https://www.linkedin.com/in/kunal-deshmukh-a94b69179)

## PROFILE

**Data Engineer with 4 years of building experience in large-scale data pipelines, ETL processes, and data warehouse solution. Technologies used such as Python, SQL, Pyspark, Databricks, Synapse, Kafka, ADF to develop multi-terabyte data solution for lifescience domain.**

## TECHNICAL SKILLS

**Programming Languages:** Python, SQL

**Big Data Technologies:** Spark, PySpark, Spark SQL

**Streaming Technology:** Kafka, Azure Event Hub, Spark structured streaming

**Cloud computing:** ADF, Databricks, ADLS, Amazon S3, Synapse

**Orchestration:** Airflow

**Familiar:** Fast Api

## EXPERIENCE

### Infosys Limit

Dec 2024 – Present

*Senior Associate Consultant*

*Bangalore, Karnataka*

- Developed a real-time data ingestion pipeline using Azure Event Hub to stream high-volume event data.
- Implemented Python-based consumers to read and process events from Event Hub efficiently.
- Integrated with PostgreSQL for intermediate staging and lookup operations during event processing.
- Loaded processed data into Delta Lake tables for scalable and ACID-compliant storage and querying.
- Optimized event processing for high throughput and low latency in a production environment.
- Implemented File, DataBase comparison system to efficiently maintain data consistency between 2 systems.

### Tata Consultancy Services

Oct 2022 – Dec 2024

*System Engineer*

*Bangalore, Karnataka*

- Expertly engineered and optimized ETL processes from various sources, ensuring seamless data consolidation, cost efficiency, and strict adherence to data standardization.
- Designed and orchestrated third-party vendor file pipelines into GDNA using Apache Airflow, optimizing data ingestion workflows, and ensuring smooth loading into Amazon Redshift for real-time analytics.
- Conducted performance tuning on SQL queries, optimizing data retrieval by 20%.
- Developed an Excel-based solution for generating outbound data for table operations, streamlining the data export process, and enhancing data handling efficiency.
- Upgraded Databricks Runtime to 13.3, resolving complex ETL issues, and created automated ETL processes to minimize manual efforts for future upgrades.

### Tata Consultancy Services

Aug 2021 – Oct 2022

*Assistant System Engineer*

*Bangalore, Karnataka*

- Implemented data quality checks, resulting in a 15% improvement in data accuracy.
- Upgraded Airflow to version 2.7.3, implemented disaster recovery protocols, and automated instance creation with OneCloud API, ensuring continuous production server uptime.

## CERTIFICATION AWARDS

**Infosys Certified Big Data Engineer**

**Fresco Play MileStone Award (TCS)**

**Elevate Wings Award (TCS)**

## EDUCATION

**RIT**

*Bachelor of Computer Science*

Goa, India

Aug. 2017 – May 2021

**MHSS**

*Senior Secondary school*

Goa, India

Aug. 2015 – May 2017