

Hammad Zaki | Data Engineer

Ranchi, Jharkhand | hammadzaki1947@gmail.com | +917762911961 | [Portfolio](#) | [linkedin.com/in/hammadzaki](https://www.linkedin.com/in/hammadzaki)

Data Engineer with **4 years of experience** in **Python, SQL**, and databases. Skilled in building scalable **data pipelines**, **web scraping** solutions, workflow automation, and developing **ETL/ELT pipelines** to improve data accuracy and efficiency. Proficient in **Databricks, Airflow, PySpark**, with strong exposure to cloud platforms like **AWS** and **Azure**.

SKILLS

- **Programming languages** : Python (PySpark, Pandas, Seaborn), SQL
- **Tools & Technologies** : Databricks, Airflow, DBT, AWS (s3, Glue, Redshift)
- **Databases & Warehouse** : Data Modeling, Snowflake, NoSQL (MongoDB)
- **AI & Machine Learning** : LLM, Langchain, NLP, Scikit-learn

WORK EXPERIENCE

Dentistry Automation - Consultant

Remote

Data Engineer

08/2025 - 02/2026

- Re-architected SQL-based ETL system into a scalable Delta Lake architecture on Databricks.
- Cloned and migrated 10+ SQL Server databases (across 1 TB data) into a unified Delta Lake architecture on Databricks.
- Designed and built a fully metadata-driven ETL framework capable of processing 140+ heterogeneous tables using a single dynamic pipeline.
- Implemented dynamic schema resolution and column mapping using configuration tables, eliminating the need for table-specific ETL code.
- Converted 50+ SQL stored procedures into optimized Spark-based Databricks jobs, standardizing execution, retries, and error handling.

BitsCrunch

Remote

Data Engineer

07/2022 - 8/2025

- Developed a **Databricks** pipeline using **PySpark** and **Web3** to fetch and process the total supply of NFT collections across more than **5 blockchains**.
- Built an **incremental DBT model** (SCD Type 1) to generate NFT rarity scores inserting data into **Snowflake**.
- Designed a **DBT snapshot model** (SCD Type 2) to capture historical NFT sales price changes and store them in **Snowflake**, ensuring a complete record of price movements for **trend and pricing analysis**.
- Designed and managed large-scale **Delta Lake** tables stored on **S3**, partitioned by blockchain and contract address, supporting over **5 blockchains and 250,000+ NFT** collections per blockchain.
- Configured an **Airflow DAG** to run **every 5 minutes** to monitor NFT prices of different blockchains.
- Created and automated a daily **ETL pipeline** and **selenium** bot to **scrape**, process, and store **10k+ NFT-related** tweets from over 100 hashtags, 1,000 users, and various keywords into a database.
- Designed charts and dashboards in **Apache Superset**, enhancing user experience and analysis efficiency by **30%**.

EDUCATION

M.Sc (Artificial Intelligence and Data Science)

Symbiosis International University

Online

06/2023 - 12/2025

Bachelor of Science (B.Sc)

Vinoba Bhave University

Jharkhand, INDIA

07/2017 - 11/2020

Core: Physics, General Elective: Mathematics

CERTIFICATES

[Data Engineering Bootcamp](#), [Practical Data Science on AWS](#), [Natural Language Processing Specialization](#),
[IBM AI Engineering](#), [IBM Data Science](#), [Google Data Analytics](#)