

TUSHAR BASAK

[Website](#) | [LinkedIn](#) | [GitHub](#) | [HackerRank](#)

Location: Kanpur, U.P, India

Email: tushar.basak@yahoo.com | Mobile: [8299312126](tel:8299312126)

DATA ENGINEER

Highly motivated and detail-oriented Data Engineer with a strong technical background and programming skills, seeking a challenging position in **data engineering**. Dedicated to designing, building, and maintaining efficient data processing systems and pipelines to enable data-driven decision making. Committed to delivering high-quality solutions and collaborating effectively with cross-functional teams.

EXPERIENCE

MedTourEasy

Data Engineer Trainee

May 2023 – May 2023

- Developed **content-based recommendation system** for cosmetics.
- Enabled **personalized product recommendations** based on chemical compositions.

EDUCATION

Manipal University Jaipur

Master of Computer Application

Jaipur, Rajasthan, India

May 2021 – June 2023

- I have done Seminar on the topic **In-Memory Data Management**.
- I have completed a specialization in both **Open Source DB Systems** and **Cloud DB System**.
- I achieved a score of **80%** without any backlogs.

PROJECTS

Real-Time Stock Market Data Processing Engine

Manipal University Jaipur

[Source Code](#)

Apr 2023 – June 2023

- Developed a real-time data processing and analysis project using Apache Kafka and AWS services, focusing on extracting and transforming data from a pre-recorded dataset.
- Implemented a Python-based mechanism to extract data from a CSV file and established a continuous stream of data using Apache Kafka, ensuring a consistent inflow of records for real-time analysis.
- Utilized AWS services such as Amazon S3, AWS Glue Crawler, and Athena for efficient data storage, cataloging, and querying, enabling seamless exploration and analysis of the processed data.

Analysis of Chemical Components

MedTourEasy

[Source Code](#)

May 2023 – June 2023

- Preprocessed and analyzed dataset of 1472 cosmetics from Sephora.
- Utilized word embedding and t-SNE for dimensionality reduction.
- Created interactive visualization using Bokeh.
- To implement efficient data management and retrieval techniques.

TECHNICAL SKILLS

Languages : Python, and SQL

Frameworks : Apache Hadoop, Apache Spark, and Apache Flink

Libraries : Pandas, NumPy, Matplotlib, PyTorch and TensorFlow

Databases : MySQL, PostgreSQL, NoSQL, MongoDB, and Cassandra

Dev Tools : Git, Jenkins, Docker, and Kubernetes