

Dhanush Shet KP

+91-8105144193 | dhanush24shet@gmail.com | [linkedin.com/in/dhanush-shet-kp/](https://www.linkedin.com/in/dhanush-shet-kp/) | github.com/Dhanushshet

PROFILE

Data Engineer with 3+ years of experience in data engineering, specializing in building efficient pipelines, optimizing SQL queries, and ensuring data quality. Proficient in Spark technologies, with a strong focus on migrating Hive to Spark for improved performance and scalability.

TECHNICAL SKILLS

Languages: Python, SQL
Big Data technologies: Spark, PySpark, Hadoop, Hive
Streaming technologies: Spark Streaming and Kafka
Data Engineering tools: Data Modelling, ETL design
Developer Tools: Git, Google Cloud Platform, VS Code
Libraries: pandas, NumPy, Matplotlib

EXPERIENCE

Nineleaps Technology Solutions - Software Developer August 2021 – Present
Client - Uber Cloudlake Platform Bengaluru, KA
Senior Data Engineer September 2024 – Present

- Delivered dynamic Tableau dashboards, improving reporting efficiency through enhanced visualization and data storytelling.
- Worked on fixing and optimizing alerts for migrated cloud pipelines, ensuring consistency, proper partitioning, and enhanced system performance.
- Migrated on-premises ETL workflows to cloud services, utilizing tools to improve scalability and ensure uncompromised data quality.
- Monitored cloud pipeline alerts and UDQ results to ensure seamless execution and validation.

Client - Uber Finance Engineering July 2023 – August 2024
Data Engineer

- Led the transformation of Hive processes into Spark 3.0 workflows, employing HDFS storage to cut data processing time by 40% and reduce costs by 60%.
- Achieved top-tier data validation by designing comprehensive testing strategies that ensured 99%+ data quality compliance across batch and distributed systems.
- Authored and implemented 100+ rigorous test cases, guaranteeing smooth deployment of workflows into production with Airflow-managed scheduling.
- Constructed scalable ETL pipelines using Hive, Spark, Kafka, and HDFS, tailored for high-throughput distributed processing tasks.

Client - Uber Customer Obsession Data Analytics August 2021 – July 2023
Junior Data Engineer

- Overhauled 70+ Tier 1 and Tier 2 pipelines handling 10B+ daily events and supporting 1M+ queries, leveraging Hive and HDFS to improve customer satisfaction by 30%.
- Engineered ETL workflows for large-scale datasets, utilizing Spark and Hadoop, achieving a 40% pipeline runtime reduction and 25% resource savings.
- Collaborated on data analysis using Airflow to improve workflows, enhancing performance and reducing errors.
- Built a reliable cloud infrastructure with GCP (BigQuery, Cloud Functions) and Kafka, ensuring 99% uptime and cutting costs by 50%.

EDUCATION

Manipal University Aug 2019 – July 2021

Canara Engineering College Aug 2015 – July 2019

