Maher Venkata Sai Srinivas | Data Engineer

| +1 (617)-545-5075 | mahervenkatasaisrinivas@gmail.com | LinkedIn |

PROFESSIONAL SUMMARY

Results-driven Senior Data Engineer with **4+ years** of experience in designing and implementing scalable data pipelines, cloud-based analytics solutions, and **high-performance ETL processes.** Proficient in **Python, SQL, Spark, and cloud platforms (AWS, Azure)** with expertise in data warehousing, performance optimization, and **machine learning** integration to drive business intelligence and decision-making.

TECHNICAL SKILLS

- **Programming Languages**: Python, SQL, Scala, Java
- **Big Data & Processing**: Apache Spark, Apache Kafka, AWS Glue, Apache Airflow
- **Databases & Data Warehousing**: PostgreSQL, MySQL, Snowflake, AWS Redshift, Delta Lake
- **Cloud & DevOps**: AWS (EC2, S3, Lambda, Glue, Redshift, CloudWatch), Kubernetes, Docker, Terraform
- **ETL & Data Integration**: Apache Kafka, Apache Airflow, DBT, Apache NiFi
- **Visualization & Reporting**: Power BI, Tableau, MS Excel
- **Data Governance & Security**: GDPR, HIPAA, SOC 2 compliance, Encryption, Access Control

PROFESSIONAL EXPERIENCE

FedEx | Senior Data Engineer   Jan 2023 – Present | USA

- Designed ETL pipelines for structured and unstructured data, improving processing efficiency by 40% and reducing latency by 30%.
- Built scalable data architectures with Apache Spark, AWS Glue, and Kafka, reducing execution time by 30% and improving throughput by 25%.
- Migrated on-premise data warehouses to AWS Redshift and Snowflake, cutting operational costs by 30% and enhancing data accessibility and performance.
- Automated 20+ data pipelines with Apache Airflow, Kubernetes, and Terraform, reducing manual reporting by 60% and increasing pipeline reliability by 50%.
- Optimized complex SQL queries for large-scale data extraction, reducing execution time by 50% and improving performance for analysts.
- Developed real-time monitoring and alerting solutions with Prometheus, AWS CloudWatch, and Datadog, ensuring 99% uptime and quick issue resolution.
- Implemented data governance frameworks to ensure GDPR, HIPAA, and SOC 2 compliance, achieving 100% audit compliance.
- Standardized CI/CD workflows with GitHub Actions, Docker, and Kubernetes, reducing deployment failures by 40% and accelerating release cycles by 30%.

Capgemini | Data Engineer   Jan 2020 – Mar 2022 | India

- Designed batch and real-time data pipelines with Apache Spark, Kafka, and AWS Glue, improving processing efficiency by 35% and reducing latency by 25%.
- Developed high-performance data models and schemas in PostgreSQL, Redshift, and Snowflake, reducing query execution time by 40%.
- Optimized complex database queries and ETL workflows, increasing transformation speed by 50%, reducing compute costs by 20%.
- Integrated cloud-based data lakes and warehouses with AWS S3, Delta Lake, and Snowflake, reducing storage costs by 25% while improving query performance.
- Automated data pipeline orchestration with Apache Airflow and Prefect, reducing manual intervention by 60% and increasing workflow reliability by 40%.
- Established CI/CD pipelines with GitHub Actions, Docker, and Terraform, reducing deployment time by 35%.

EDUCATION

Sacred Heart University   April 2022 – June 2023 | Fairfield, Connecticut
Master of Science in Business Analytics

K L University   August 2018 – June 2021 | India
Bachelor of Business in Business Administration

PROJECT

Optimizing Data Pipelines for E-Commerce Analytics

- Built an ETL pipeline using Apache Spark and AWS Glue, processing 100M+ records daily, Reduced data latency by 60% via real-time Kafka streaming, Enhanced query performance with data partitioning and indexing.
- **Skills:** ETL (Extract, Transform, Load), Data Pipeline Optimization, Real-Time Data Processing, Data Warehousing & Big Data Management, Performance Optimization (Indexing, Partitioning, Caching), Data Visualization & Business Intelligence, Cloud Computing & Scalability

Customer Segmentation for Personalized Marketing

- Engineered data pipelines for customer segmentation analysis using Python and SQL, applied clustering techniques to group customers based on purchasing behavior, delivered insights to the marketing team, improving targeted campaign effectiveness.
- **Skills:** Machine Learning (Clustering, RFM Analysis), Customer Behavior Analysis, Personalization & Recommendation Systems, A/B Testing & Experimentation, Data Preprocessing & Feature Engineering, Marketing Analytics

IoT Sensor Data Processing for Smart Cities

- Designed a real-time data processing framework for IoT sensor streams using AWS Lambda.
- **Skills:** IoT Data Collection & Processing, Real-Time Analytics & Edge Computing, Time-Series Data Analysis, AI & Predictive Maintenance, Cloud Computing & Distributed Systems, Security & Authentication in IoT