

LEVI GAGNE

Data Engineer

PROFESSIONAL SUMMARY

Data Engineer focused on optimizing ETL processes and delivering scalable reporting solutions. Experienced in Python, PySpark, and SQL, with expertise in AWS, Hadoop, and Databricks for developing high-performance, cloud-native data pipelines. Skilled in quantitative analysis with a background in applied econometrics, enhancing machine learning models and AI-driven insights. Proven ability to translate complex business needs into scalable, impactful data solutions across industries, from automotive to environmental science.

EMPLOYMENT HISTORY

DATA ENGINEER

Jun 2021 - Jun 2024

General Motors

Austin, TX

- ♦ Automated end-to-end ETL workflows using PySpark on Kubernetes, integrating data from S3 and Hadoop to optimize transformations and reduce manual intervention across pipelines.
- ♦ Developed a Python-based real-time monitoring solution for production jobs, incorporating automated alerts and extended log retention to support issue resolution beyond standard retention periods.
- ♦ Enhanced Spark job performance by utilizing logical and physical execution plans, improving processing speed, scalability, and cost-efficiency for high-volume ETL tasks.
- ♦ Played a role in the creation and automation of 14 KPIs in SQL for 96 daily reports, including metrics such as daily sales and prior year month-to-date sales, by querying data from sales, calendar, and dimension tables, then transitioning them to Spark SQL for automation.
- ♦ Contributed to the migration from EC2/Kubernetes to Databricks, including the transfer of workflows, the development of new processes, and the validation of tables to ensure accuracy after the migration.
- ♦ Led weekly learning sessions for teammates on Databricks, focusing on topics like managing Spark in notebooks, mounting S3 buckets, and utilizing Spark-powered notebooks over traditional Jupyter notebooks.

GRADUATE RESEARCH ASSISTANT

Aug 2020 - May 2021

Central Michigan University

Mount Pleasant, MI

- ♦ Supported faculty in research, enhancing data accuracy and website management.
- ♦ Guided students in statistical and economic analysis, improving academic performance.
- ♦ Collaborated on environmental economics projects, contributing to innovative research models.

ENVIRONMENTAL, HEALTH, & SAFETY INTERN

Jun 2020 - Aug 2020

Amazon.com, Inc.

Seattle, WA

- ♦ Conducted risk assessments, reducing workplace incidents and improving safety standards.
- ♦ Enhanced equipment safety protocols, minimizing potential hazards in the work environment.
- ♦ Analyzed job hazards, contributing to a safer workplace with measurable safety improvements.

TOXICOLOGY & ENVIRONMENTAL RESEARCH INTERN

Jul 2016 - May 2021

Dow Chemical Co

Midland, MI

- ♦ Collected R&D samples for clinical studies, enhancing data accuracy and study reliability.
- ♦ Executed complex histological procedures, ensuring precise slide preparation for analysis.
- ♦ Prepared laboratory solutions and reagents, maintaining adherence to industry standards.

EDUCATION

MASTER OF ARTS: ECONOMICS

Aug 2019 - May 2021

Central Michigan University

Mount Pleasant, MI

SKILLS

Python, SQL, Hadoop, Apache (Airflow, Iceberg, Kafka, NiFi, Parquet, Spark),
AWS (EC2, Glue, Lamda, S3), Databricks, ETL, Kubernetes, Linux, Unix, Stata, Econometrics,
Jupyter Notebooks, Data Warehousing, Data Modeling, Data Lake Management, Data Validation.