AUSTIN, TX • LEVIGAGNE0@GMAIL.COM • (989) 423-7334

LEVI GAGNE

Data Engineer

PROFESSIONAL SUMMARY

Data Engineer specializing in Python, Spark, and SQL, with expertise in AWS, Hadoop, and Databricks for building high-performance, cloud-native data pipelines. Skilled in optimizing ETL workflows and delivering scalable data products. Strong background in quantitative analysis and applied econometrics, with a history of delivering data science and reporting products into production.

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.
- 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 ASSISTANT
Aug 2020 - May 2021
Central Michigan University — *Mount Pleasant, MI*

- Graded assignments, entered scores, and occasionally lectured, providing administrative and instructional support to ensure smooth course operations.
- Guided students in statistical and economic techniques, enhancing their understanding through tailored support in lectures, assignments, and one-on-one sessions.
- Created a digital archive of economic research from historical newspapers, developing a searchable filing system by categorizing topics and years from manually processed records.

ENVIRONMENTAL, HEALTH, AND SAFETY (EHS) INTERN
Jun 2020 - Aug 2020
Amazon.com, Inc. — *Seattle, WA*

- Assisted with risk assessments by identifying workplace vulnerabilities and learning best practices for reducing incidents and improving safety standards.
- Observed and supported the development of equipment safety protocols, gaining hands-on experience in identifying and mitigating potential hazards.
- Participated in job hazard analyses, contributing to discussions on workplace safety improvements while building foundational knowledge in risk management strategies.

PATHOLOGY CO-OP
Jul 2016 - May 2021
Dow Chemical Co — *Midland, MI*

- Collaborated with toxicologists and pathologists to collect R&D samples, supporting regulatory requirements and internal product research.
- Helped process histological samples for slide analysis, preparing high-quality slides used in studies for regulatory submission and internal evaluations.
- Learned and assisted in various aspects of toxicology research from study start to finish, gaining the skills needed to eventually conduct studies independently.

EDUCATION

MASTER OF ARTS: ECONOMICS   **Aug 2019 - May 2021**
Central Michigan University   *Mount Pleasant, MI*

BACHELOR OF SCIENCE: ECONOMICS & ENVIRONMENTAL   **Aug 2018 - May 2020**
STUDIES
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.