

Brian Sterle

sterle.brian@gmail.com ❖ (937) 620-0860 ❖ Dayton, OH ❖ www.briansterle.com

WORK EXPERIENCE

General Motors

June 2016 – July 2023

Senior Data Engineer | Jan. 2021 – July 2023

Austin, TX

Led the design, architecture and implementation of Big Data Services teams' many apps and services.

- Created the application responsible for ingesting, normalizing and refining all of GM's data into its centralized data lake and data warehouse. App processes 3 trillion records and 300+ terabytes daily.
 - Cloud-native based implementation uses Scala, Spark, Delta Lake, Docker, Kubernetes and gRPC (replaced a legacy application using Java, MapReduce, and CSV).
- Enhanced data pipeline efficiency and job scheduling, significantly reducing resource allocation time and enabling the processing of millions of jobs, including 5.8 million base ingestion jobs.
- Spearheaded a major data engineering effort to build out a modern data lake & cluster, migrating thousands of business-critical ETL jobs from on-prem YARN Hadoop Cluster to on-prem Kubernetes with Spark
- Interacted with users to answer questions and integrate their feedback, and mentored a team of 6 junior data engineers, providing guidance on best practices in data processing & architecture, pipeline optimization, and software development.

Data Engineer | June 2016 – Dec. 2020

- Implemented a custom streaming solution to ingest real-time media usage data from vehicles and merged it with geospatial radio data enabling licensing to a major satellite radio company on a multimillion-dollar contract.
- Developed a web-based data transfer application, now an integral component of the data teams' tool set, facilitating secure and user-friendly transfers between production and pre-production environments.
 - Implemented with Spring Boot backend and Angular frontend, the app handled over 47,000 jobs, copying approximately 2 million GB of data and serving nearly 1,900 users.
- Created a Spark-based application which enabled daily data loads to HDFS from RDBMS systems such as Oracle, SQL Server and Postgres, and re-synchronized tables normally loaded incrementally.
- Programmed a data workspace management tool which boosted collaboration and productivity among data professionals.
 - Managed thousands of user and data requests, including over 26,000 user additions and 23,000 table additions, significantly enhancing data project setups.

Huntington Bank

May 2015 – Aug. 2015

Data Analytics Intern

Columbus, OH

- Developed a comprehensive debit card fraud detection system using R and integrating various data sources, resulting in actionable insights for staffing and loss mitigation.

EDUCATION

Ohio State University

Aug. 2012 - June 2016

B.S., Major in Industrial & Systems Engineering, Minor in Computer Science

Columbus, OH

- *cum laude*; 3.6/4.0 GPA

SKILLS & INTERESTS

- **Skills:** Scala; Java; Go; Rust; Python; R; Spark; Delta Lake; Kafka; Hadoop/HDFS; Kubernetes; Docker; Spring Boot; Linux; TypeScript/JavaScript; Angular; Oracle; S3; Postgres; SQL Server
- **Interests:** philosophy; bodybuilding; chess; theology; dogs; bitcoin; pc gaming; music