

Mason Daniel

832-704-5932 | masondanielt@gmail.com | [LinkedIn](#) | [Github](#)

SUMMARY

Motivated software developer with a passion for simple solutions and readable code. Thrives on working with others and enjoys learning new technologies.

EXPERIENCE

Big Data Software Developer

July 2020 – Present

General Motors

Austin, TX

- Became lead developer of a data pipeline process that batch loaded data from HDFS to a PostgreSQL database, enabling over 100 million records loaded per day with 99.9% SLA attainment
- Designed and developed a Java application that cleaned 5000+ tables of data per day via Apache Spark
- Created a Java Spring Boot microservice APIs that received over 1 million http requests per day
- Converted 2 Java applications and 3 Java Spring APIs to a Docker & Kubernetes environment as part of the next generation platform for internal company clients, resulting in 24/7 uptime
- Appointed as the point of contact for the pipeline process and took ownership of all development within 7 codebases
- Used SQL to directly manage over 200,000 database records of metadata

Full Stack Developer

May 2020 – August 2020

Southwestern University

Georgetown, TX

- Hired by the University's Study Abroad Director to clean datasets of potential study abroad programs, and build a web application allowing students to view them via filter-based search
- Cleaned over 20000 rows of data using Apache Spark
- Implemented the front-end search page using React and the back-end database using node.js and PostgreSQL
- Held regular meetings with the client to adapt the product to their needs
- Shipped the application July 2020, which served as a prototype for a later implementation utilizing the University's website CMS

SKILLS

Java, JavaScript, Python, Hadoop/HDFS, Apache Spark, Kubernetes, Docker, Oracle/MySQL, PostgreSQL, JDBC, Spring Boot, PCF, AWS, Hive, Linux, node.js, React, Git

PROJECTS

DataLake with AWS | Python, Apache Spark, AWS, S3

- Built functional python ETL script with functions that initialized spark clusters using pyspark library to extract songs stored in S3 bucket
- Partitioned songs data by year and artist and compressed in parquet output files to increase load performance
- Used the overwrite mode in spark to ensure every new run of ELT script is overwritten in data lake to avoid duplicates

EDUCATION

Southwestern University

Bachelor of Arts in Computer Science

Georgetown, TX

Aug. 2016 – May 2020