Drew Gustafson

Seattle, WA • drewgustafson12@gmail.com • (206) 849-2709 • Portfolio: dgust.cc • linkedin.com/in/dtg3

PROFESSIONAL SUMMARY

Senior Data Engineer with 4+ years of experience designing metadata-driven ETL systems and building data warehouses using Azure. Led the design of an ETL framework at WTW and delivered scalable data solutions at Insight. Skilled in integrating disparate data sources, optimizing pipeline performance, and developing robust systems that reduce development effort for new projects.

WORK EXPERIENCE

Willis Towers Watson | Data Engineer | Remote | June 2023 - Present

● Designed and implemented a metadata-driven ETL framework in Azure, significantly reducing code redundancy and minimizing the number of ADF pipelines and Databricks notebooks. Supported multiple data sources including SFTP, Delta Sharing, and Event Hub. Documented framework and trained teammates on its usage.
● Designed the framework with scalability in mind, allowing future integration of sources such as high-volume OLTP databases on SQL Server, Salesforce, or Oracle with minimal code changes.
● Replaced a legacy fixed-sequence scheduling system, which relied on manually assigned job order numbers, with a metadata-driven DAG architecture using dependsOnJobKey for dynamic, dependency-aware orchestration.
● Developed dynamic SFTP scheduling that detects file arrivals and queues workflows, eliminating fixed scheduling and improving reporting SLAs.
● Optimized incremental loads by fixing boundary logic bugs and switching SFTP ingestion from fixed lookback periods to watermark-based incremental loads using ETL-derived timestamps.
● Increased pipeline resilience by developing Python UDFs to handle CSV irregularities, including embedded commas, line breaks, inconsistent delimiters, and schema drift.
● Built a pipeline monitoring tool using Python and Azure Logic Apps that analyzes logs to report workflow status, row counts, and last update times, enabling proactive detection of silent failures and stale pipelines.

Insight Enterprises (Clients: Microsoft, Gilead, Microchip, TQL) | Remote | April 2021 - May 2023
Career Progression: Associate Data Engineer - Data Engineer - Senior Data Engineer

● Built Azure-based data warehouses by ingesting millions of rows from sources such as SQL, Salesforce, CSVs, APIs, and web scraping. Transformed data through a medallion datalake architecture using ADF, SQL, and PySpark.
● Designed and delivered a full-stack data platform, modeling denormalized flat files into a normalized, Kimball-style star-schema warehouse, and developed Power BI reports tailored to client needs.
● Engineered parameterized, metadata-driven PySpark notebooks to refresh hundreds of cloud tables.
● Led a Synapse migration project, implemented scheduling, job configurations, and incremental loads via ADF and stored procedures on the dedicated SQL pool.
● Deployed Azure infrastructure for a greenfield data platform with Terraform; leveraged and maintained existing CI/CD pipelines for deployment automation.
● Supported ML models by building data pipelines for product recommendations and defect detection using Databricks and CNNs, including a patent-pending metadata-driven framework.
● Built Python-based validation tools for row-level table diffs and duplicate key detection across environments.

USI Insurance Services | Operations Data Analyst | Remote | March 2020 - April 2021

● Automated reports and data prep tasks using PowerQuery and Excel, sparking transition into data engineering.

EDUCATION

University of Washington, Michael G. Foster School of Business | Seattle, WA | September 2016 – June 2021
M.S. Information Systems (GPA 3.87), B.A. Business Administration - Info Systems & Supply Chain (GPA 3.62)

TECHNICAL SKILLS

Languages: Python, PySpark, Pandas, SQL (Recursive CTE, MERGE, DAG, Window Functions), Terraform
Azure: Data Factory, Databricks, Synapse, Blob Storage, Key Vault, DevOps, Event Hub, Logic Apps
Other: Data Modeling, Lakehouse Architecture, Power BI, Tableau, ML Support, CI/CD, Git