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 owning production-grade, scalable batch and streaming data pipelines using Azure. Designed metadata-driven orchestration frameworks and delivered real-time file detection and alerting systems. Experienced in supporting analytics and ML workflows through robust, high-performance data systems.

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.
● Modernized scheduling architecture by implementing metadata-driven DAGs, enabling dependency-based orchestration and eliminating brittle fixed-order sequencing.
● Built a streaming-like ingestion mechanism that detects file arrivals from SFTP and dynamically orchestrates downstream workflows, improving SLAs and reducing latency in reporting pipelines.
● 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.
● Designed and owned an automated monitoring tool using Python and Logic Apps to track pipeline health, identify silent failures, and surface stale data, improving observability and operational reliability.

Insight Enterprises (Clients: Microsoft, Gilead, Microchip, TQL) | Remote | April 2021 - May 2023
Career Progression: Associate Data Engineer - Data Engineer - Senior Data Engineer
● Ingested and transformed millions of rows across batch pipelines using ADF, Databricks, and SQL; supported high-scale data modeling in a medallion architecture for multiple client-facing dashboards.
● 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.
● Built data pipelines powering ML models for product recommendations and defect detection, collaborating with data scientists to deliver production-ready feature inputs and refresh logic.
● 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