

# Melissa Mason

Saline, MI 48176

masonlm68@gmail.com | (734) 294-2766

GitHub: [github.com/mel1441](https://github.com/mel1441)

Website: [www.mmason.tech](http://www.mmason.tech)

---

## Senior Data Engineer

Experienced architect of reproducible, solution-oriented data infrastructures across municipal, geospatial, and public sector domains. Expert in Python, PySpark, and AWS-native ETL pipelines with a focus on data quality, rollback hygiene, and semantic traceability. Skilled in mentoring teams, modernizing legacy systems, and delivering scalable pipelines that power internal analytics and client-facing applications.

---

## Work Experience

### Senior Data Engineer — Munetrix (Remote)

April 2022 – September 2023

- Designed and deployed data completeness reporting, enabling stakeholders to assess product readiness in real time.
- Led migration from manual to automated ETL pipelines, improving data delivery speed and reducing manual QA effort.
- Maintained PostgreSQL warehouse and AWS ETL stack (Glue, Lambda, Athena) for municipal intelligence platform.
- Identified data quality issues and led team members through corrective actions.
- Audited MySQL warehouse post-merger to streamline education product data footprint.
- Mentored new engineers, accelerating onboarding post-acquisition.

### Data Engineer — Dynamo Metrics (Remote)

August 2016 – April 2022

- Architected and maintained core SaaS product across frontend and backend, supporting municipal analytics and geospatial insights.
  - Designed and optimized PostgreSQL data warehouse and geospatial architecture using PostGIS and QGIS.
  - Reduced data processing time from days to hours by rewriting and optimizing transformation queries.
  - Migrated legacy on-prem warehouse to AWS, improving scalability and reducing infrastructure costs.
  - Built automated ETL pipelines using AWS Glue and PySpark, cutting transformation latency and operational overhead.
  - Delivered data engineering support for projects with the Michigan Economic Development Corporation and NIH, ensuring reproducibility and auditability.
  - Developed proprietary algorithms to classify tax and sale statuses from time series property data—an industry-first innovation.
-

## **Skills & Technical Proficiencies**

### **Data Engineering & Architecture**

- ETL pipeline design (Event-driven, batch, streaming, historical)
- Data orchestration (Dagster, AWS Step Functions, Apache Airflow)
- Data architecture design (Medallion, Mesh, Fabric, lake, warehouse)
- Geospatial data modeling and analysis
- Data auditing, quality assurance, and semantic traceability
- Schema design (Star, Snowflake) and validation across structured (CSV, SQL), semi-structured (JSON, XML), and geospatial (GeoJSON, shapefiles) formats
- Query optimization, data transformation, and reproducibility tooling
- Python-based API clients, requests, and SDK integration
- JSON payload manipulation, schema validation, and error handling
- API-driven ETL orchestration and cloud-native data movement
- Manual to automated data migration

### **Cloud & Infrastructure**

- AWS (Glue, Lambda, Step Functions, S3, Athena, RDS)
- GCP (BigQuery, Compute Engine, Cloud Storage)
- Infrastructure as Code (IaC) (Terraform), Docker, GitHub Actions, version control (Git/GitHub)

### **Programming & Tooling**

- Languages: Python, PySpark, SQL, T-SQL, PL/SQL, TypeScript, Bash, Powershell, HTML, CSS
- Libraries: Pandas, NumPy, Requests, SQLAlchemy, Boto3, GeoPandas, pdfplumber, pytesseract, BeautifulSoup
- Databases: PostgreSQL, MySQL, SQLite, MSSQL, DuckDB
- Frameworks: React, FastAPI
- Applications: DBeaver, pgAdmin, QGIS, Jupyter Notebook, VSCode, Tableau

### **Collaboration & Delivery**

- Agile/Scrum methodologies, project management (Jira, Slack)
- Cross-functional mentoring and post-merger onboarding
- Data documentation (Confluence), stakeholder reporting, and product readiness metrics

---

## **Education**

### **University of Michigan**

Bachelor of Arts, Political Science — 2016