Christopher MacMaster

510-334-0387 | christopher.macmaster@gmail.com Portfolio | GitHub | LinkedIn

CAREER SUMMARY

Data Engineer with 3 years of experience building ETL pipelines and data transformation logic for election intelligence platforms. Proficient in Python (Pandas, Polars, regex), AWS (S3), PostgreSQL, Google Colab, and orchestration tools (Apache Airflow, Dagster). Passionate about data integrity and civic transparency, with experience delivering client-facing features for county election officials.

SKILLS

Languages: Python, JavaScript, SQL, HTML, CSS
Databases: PostgreSQL
Cloud: AWS (S3)
Data & Orchestration: Polars, Pandas, Dagster, Apache Airflow
Other: Google Colab, regex, FileZilla, Sequel Ace

EDUCATION

Harvard College | B.A. | GPA: 3.78 August 2015 - May 2019
Relevant Coursework: Survey Research Methods, American Public Opinion, Modern American Political Campaigns, Progressive Alternatives.

App Academy October 2022 - May 2023

- Six-month full-stack web development bootcamp. Capstone project was GroupUp, an Overwatch-themed social media application modeled after the site Meetup.

EXPERIENCE

Data Engineer - *Civera* November 2023 - October 2026
Cambridge, MA

- Joined Civera to take over and streamline the data ingestion process for the ElectionStats platform, developing Python-based transformation logic in Google Colab using Pandas to clean and denormalize county voting data from diverse sources (Excel files, XMLs, PNGs, Dominion/Hart/Unisyn vendor exports), delivering standardized CSVs to a legacy importer.
- Reduced notebook transformation runtime by refactoring file I/O patterns and introducing intermediate data extracts, eliminating redundant re-reads and enabling faster iterative development.
- Built extensive custom logic on top of Tabula and Camelot for edge-case-heavy formats like Sonoma County, CA, to extract table data from imperfect PDFs that the libraries read inconsistently out of the box.
- Worked with a senior engineer to migrate value extraction from string helper functions to compiled regex with capture groups, then owned the regex work on the data itself — parsing raw contest name strings and assigning structured variables (Office, DivisionName, etc.) based on string content, improving maintainability across counties and election events.
- Collaborated on the migration from notebook-based Pandas processing to an Airflow and eventually Dagster production pipeline leveraging Polars, contributing domain expertise to implement faster, memory-efficient transformations that reduced processing time and improved scalability for county and statewide clients.
- Engineered an adjudication flagging feature for Civera's BallotVerifier platform, developing a graph-backed asset within Dagster to identify and label adjudicated ballots. Enabled Ada County, Idaho, to filter, count, and display ballots separately – delivered ahead of deadline with zero defects.
- Onboarded and mentored multiple data engineers as the team scaled, providing hands-on training in transformation logic, county-specific data quirks, regex patterns, and debugging techniques.
- Acted as primary technical point-of-contact for project managers and downstream stakeholders, translating county-specific client requests into technical solutions and ensuring data accuracy.
- More recently, leveraged AI-assisted development (Claude) to accelerate output — directing higher-level design, testing plans, and method selection while using AI to implement fixes, catch front-end bugs, and build new data-quality checks (missing images, missing CVRs) that surface client errors days earlier.