Aidan Draper

(Software) Data Engineer

✉ adraper19@gmail.com | 🏠 aidandraper.com | ○ adraper2 | ◈ adraper19 | in aidan-draper

Experience

WeWork   *New York, NY*

Data Engineer   *June 2022 - Present*

- **Architecting an Airflow CI/CD workflow in Github Actions** that syncs to S3, builds a Docker image for ECR and deploys to a private EC2 server.
- Designing modern data team processes, such as dbt SQL design and DAG testing standards. **Integrating linting, code security (safety & bandit) checks, and PEP8 formatters** into the CI workflow. Backfilling unit testing on DAGs and setup PR guidelines.
- **Augmenting a reporting layer** (Tableau/Snowflake/Tray.io) **to better suit the needs of our non-technical stakeholder groups.** Increasing data literacy by training out on Tableau, Snowflake, and data modelling best practices.
- **Wrapping up a sunset of an AWS RDS instance and migrating DAGs to load into a data warehouse in Snowflake** to simplify the data lifecycle and save thousands of dollars on hosting duplicate data.
- Managing Airflow to S3 DAGs that stage CSVs to be transformed in dbt or ingested with Fivetran connectors from sources such as Workday, Greenhouse, and Qualtrics as well as internal apps and data science models.

Pathstream *(Series A Startup)*   *New York, NY*

Data Engineer   *Nov 2021 - May 2022*

- Worked in a team of 4 to **deploy a custom OLAP and REST client in Django in just 3 months**. It handles tasks such as data warehousing, table denormalization, and syncing internal data with ingested external data. See **https://gitlab.com/frame-00/00-api**.
- **Wrote ETL pipelines from scratch in Python using Airflow** from sources such as external marketing data from TikTok and internal student data from an in-house Rails application. Revised existing pipeline tests to improve coverage.
- Migrated Airflow operators to send data to a warehousing client. Created custom queries in a Postgres DB with Django dunder methods.
- **Created and maintained SLAs by meeting with non-technical stakeholders**, and monitoring logs on DataDog and Airflow.
- Authored & reviewed technical content for data analytics courses as well as documentation for our data engineering repositories.

Teaching Assistant - Data Analytics (Contract)   *June 2021 - Dec 2021*

- The role included answering technical questions in office hours and emails about topics in Sheets, Tableau, SQL, Statistics, and Python courses.

University of Notre Dame   *Notre Dame, IN*

Computer Vision Research Assistant (PhD Student)   *Jun 2019 - June 2021*

- Investigated using transfer-learned multiview scene rendering neural networks (NN) for distortion correction modelling on 3D fingerprint images to improve minutiae map matching accuracy with traditional 2D fingerprints images. **Trained self-supervised NNs on >3 million frames** from video. **Used high throughput computing to parallelize code across 20 GPUs.**

Summer Research Assistant   *Jun - Aug 2018*

- **Built a classifier to measure grass species abundance at 85% accuracy** with Landsat satellite images processed in R.
- Reconstructed a Bayesian hierarchical model that used a Poisson point process and parallelized it to create an **execution time speedup of 3x**.

Skills

Languages   Proficient: [Python, Bash, SQL, JavaScript, R] Experience: [Ruby, C++, Java]

Libraries   Airflow, dbt, Django, React, py: [Snappy/pyarrow, pytest, factory, Tensorflow, PyTorch, CV2, Pandas, NumPy], exposure to Rails

Tools   Snowflake, Fivetran, MySQL, Postgres, Git, Github Actions (CI/CD), Apache Parquet, JIRA, Notion, Tableau, Sheets Macros
      AWS: [ECS/EC2, ECR, Cloudfront, S3, IAM, Route 53]

Education

University of Notre Dame   *Notre Dame, IN*

Graduate Coursework in Computer Science, Advised by Dr. Adam Czajka, GPA: 3.7/4.0   *Jun 2019 - Jun 2021*

Elon University   *Elon, NC*

BS in Computer Science, BS in Statistics, GPA: 3.6/4.0   *Aug 2015 - May 2019*

Competitions

Manhattan College Business Analytics Competition   *New York, NY*

Best Poster Award for interpreting NYC small business contract spend and allocation data.   *May 2018*

Champion Analytics Case Competition   *Burlington, NC*

1st Place out of 20 teams in analyzing Champion's product consumer reviews.   *October 2017*