

Aidan Draper

(SOFTWARE) DATA ENGINEER

✉ adraper19@gmail.com | 🌐 aidandraper.com | 🌱 adraper2 | 🐙 adraper19 | in aidan-draper

Experience

WeWork

June 2022 - Present

DATA ENGINEER

New York, NY

- Led the decision, design and process of deprecating a single Airflow EC2 instance for a Docker and ECS microservices architecture that horizontally and vertically scales to **improve ELT task concurrency and reduce monthly server cost by roughly 30% by less resource idling**.
- Caught, corrected, and informed upstream stakeholders of critical bugs, including **the overcounting of headcount by 3% and salary run rate by a few million dollars**, due to outdated business logic in a few dbt models being queried for weekly reports sent to executives.
- **Architected and delivered a continuous integration and deployment (CI/CD) pipeline** in GitHub Actions for our Airflow + dbt code repository, while creating merge request guidelines and code documentation for team readability, onboarding, and SOX compliance.
- **Refactoring a Snowflake data mart with dbt models that better fit the current need and business logic** of the People Team divisions: Recruiting, Learning, Diversity/Inclusion, Legal, and Compensation, using internal app, Greenhouse, Workday, Salesforce and Shareworks data.
- **Upleveling current code development processes** by introducing a deployment workflow lifecycle with test coverage expectations, a PEP8 code formatter/linter, and package vulnerability checks while improving DAG and dbt test coverage as well as Slack alerts for workflow failures.

Pathstream (Series A Startup)

June 2021 - May 2022

DATA ENGINEER

New York, NY

- Collaborated in a team of 4 to **deploy a custom OLAP and REST client in Django in just 3 months** that handles tasks such as data warehousing, table denormalization, and syncing company internal data with ingested external data. See <https://gitlab.com/frame-00/00-api>.
- **Built 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.
- **Designed and maintained SLAs by meeting with non-technical stakeholders**, and monitoring logs on DataDog and Airflow.
- Authored and reviewed technical content for data analytics courses as well as documentation for our data engineering repositories.
- **Wore many hats** answering technical questions in student office hours and emails in topics like Python, SQL, Tableau, Statistics, and Sheets.

University of Notre Dame

Summer 2018; June 2019 - June 2021

COMPUTER VISION RESEARCH ASSISTANT (PhD Student)

Notre Dame, IN

- Studied the effectiveness of transfer-learned depth estimation neural networks for improving matching scores of 3D to 2D fingerprint images.
- **Distributed thousands of model training jobs across a GPU cluster** to find convergence of self-supervised neural networks.
- Implemented pyarrow to **improve memory management and processing speed of >3 million high-resolution frames** of fingerprint video.

SUMMER RESEARCH ASSISTANT

- Applied multi-core concurrency using R Snow library and wrote bash that deployed distributed jobs to a cluster, which **resulted in a 3x speedup in execution time** for the cross-validation process of a Bayesian hierarchical model **without losing accuracy in the model fit**.

Skills

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

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

Tools Snowflake, Fivetran, MySQL, Postgres, Redis, React, Git, Github Actions (CI/CD), Apache Parquet, JIRA, Notion, Tableau

AWS: [ECS, EC2, ECR, RDS, VPC, Cloudfront, S3, IAM, Parameter Store, Route 53]

Education

University of Notre Dame

Jun 2019 - Jun 2021

GRADUATE COURSEWORK IN COMPUTER SCIENCE, ADVISED BY DR. ADAM CZAJKA

Notre Dame, IN

Elon University

Aug 2015 - May 2019

BS IN COMPUTER SCIENCE, BS IN STATISTICS

Elon, NC

Competitions

Manhattan College Business Analytics Competition

May 2018

BEST POSTER AWARD FOR INTERPRETING NYC SMALL BUSINESS CONTRACT SPEND AND ALLOCATION DATA.

New York, NY

Champion Analytics Case Competition

October 2017

1ST PLACE OUT OF 20 TEAMS IN ANALYZING CHAMPION'S PRODUCT CONSUMER REVIEWS.

Burlington, NC