

# David Freitag

New York, NY | (262) 271-5365 | [dkfreitag@gmail.com](mailto:dkfreitag@gmail.com)

LinkedIn: [linkedin.com/in/davidkfreitag](https://www.linkedin.com/in/davidkfreitag)

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

Senior Data Engineer skilled in building mission-critical, reliable, efficient, high-volume data pipelines. History of effective project leadership, stakeholder management, and mentorship of junior engineers. Former high school teacher.

## EXPERIENCE

---

### Senior Data Engineer

February 2023 - Present

American Family Insurance

- Project leader: led a team of three engineers to build an ETL pipeline using Google Bigquery, Airflow, Docker, and Kubernetes to refresh AmFam's largest auto modeling dataset, reducing the time-to-market cycle by months.
- Improved automation of multiple vendor file transfer processes using Python, GCP Storage API, Google Secret Manager, and GNU Privacy Guard. Decreased person-hours per transfer from 2 hours to 5 minutes per transfer.
- Mentored a Rotation Analyst in technical skills including Python, BigQuery/SQL, Git/GitLab, Airflow, Terraform, and various services in Google Cloud Platform (GCP) and Amazon Web Services (AWS).

### Data Engineer

August 2021 - February 2023

American Family Insurance

- Led my team's Data Quality Initiative. Created best practices and guided implementation of automated processes for monitoring data quality. Automated identification of data quality issues monthly across 15+ datasets.
- Decreased runtime for an insurance re-rating process from 4 days to 1.5 hours by distributing the workload across multiple Kubernetes pods orchestrated with Airflow.
- Developed a fully automated ETL pipeline with Python, SQL, Airflow and GitLab CI/CD to produce a monthly product analysis dataset with 600+ attributes.
- Built a process for ad-hoc AWS Athena to GCP BigQuery data transfers using Airflow. Transferred terabytes of data using this process.

### Data Science Intern

March 2021 - July 2021

Pitney Bowes

- Analyzed global container scan data using Snowflake SQL, Python, and statistical testing to provide stakeholders with visibility into the Pitney Bowes network.
- Trained, tested, and tuned a machine learning model using LightGBM to predict cycle time for container journeys between locations within the Pitney Bowes network to improve container arrival forecasts.

## EDUCATION

---

### Baruch College, The City University of New York

May 2021

*Master of Science, Information Systems, Data Analytics*

GPA: 4.0

**Honors:** Beta Gamma Sigma Honor Society, Beta Alpha Psi Honor Society

## PROJECTS

---

### [Build Your First Serverless Data Engineering Project](#)

- I teach an intro-level data engineering course I designed on serverless data engineering in AWS. Technologies include: AWS Lambda, S3, Athena, Kinesis, Glue, and Grafana.