

# Jake Brehm

✉ mail@jakebrehm.com   [in linkedin.com/in/jakebrehm](https://www.linkedin.com/in/jakebrehm)   [github.com/jakebrehm](https://github.com/jakebrehm)   🏠 Atlanta, GA

## Skills

---

**Languages:** Python (*Pandas, NumPy, Matplotlib, PySpark, Requests*) • SQL • R • JavaScript • HTML • CSS • LabVIEW

**Services:** AWS (*EC2, S3, RedShift, SageMaker*) • Databricks • Tableau • Power BI

**Technologies:** Databases (*SQL Server, PostgreSQL, MySQL, SQLite*) • Docker • Git • TortoiseSVN • REST APIs

## Experience

---

### Horizon Health Services

Remote • Getzville, NY, USA

#### Data Engineer

Jun 2023 - Present

- Construct data pipelines with scalable architectures in Python and T-SQL for general business consumption.
- Communicate with stakeholders and IT to translate complex business logic into robust data solutions and infrastructure.
- Orchestrate ETL pipelines using cron jobs to support daily data refresh of 86 Power BI dashboards serving nearly 1200 users.
- Automate data extraction from spreadsheets, REST APIs, and SFTP servers into an on-premises database.
- Spearheaded internal Python codebase development and training, improving departmental efficiency and collaboration.
- Refactored over 30 T-SQL stored procedures to optimize queries and ensure near 100% data accuracy.
- Ported over 20 C# .NET API endpoints to Python, improving pipeline reliability and reducing execution time by 90%.
- Leveraged Active Directory to create a data crosswalk that links employee identification numbers across 3 systems.
- Facilitate discussions with marketing about integrating engagement data from 5 social media REST APIs for KPI computation.

### Virufy

Remote • San Francisco, CA, USA

#### Associate Data Scientist • Volunteer

Jan 2023 - Nov 2023

- Trained machine learning models on AWS SageMaker using data stored both locally and in S3 buckets.
- Generated over 100,000 synthetic audio files using GANs to mitigate a 1:3:3 class imbalance in training data.
- Achieved a 4.49 inception score, indicating the generated audio was high quality, diverse, and similar to the training data.

### Viewpoint Systems Inc.

Buffalo/Rochester, NY, USA

#### Systems Engineer

Apr 2020 - Jun 2022

- Communicated with the client to scope and quote new projects, as well as collaborate on existing projects.
- Obtained 8 LabVIEW certifications covering topics ranging from fundamentals to scalable architectures and design patterns.
- Conceived a version of automated test software that runs without hardware, simplifying the client's development cycle.
- Validated spacecraft thrust vector control system functionality programmatically and led demonstrations for stakeholders.

### Young & Franklin Inc./Tactair Fluid Controls Inc.

Liverpool, NY, USA

#### Development Engineer

Aug 2018 - Jan 2020

- Automated regular testing of a digital motor controller using Python to reduce the manual workload of a team of 4.
- Launched data analysis and visualization software, allowing for interactive collaboration with stakeholders and colleagues.
- Produced computer vision software that measures oil leakage with over 90% accuracy to meet compliance requirements.

## Projects

---

### ManyMiles

Nov 2023 - Present

- Created a full stack web application using *Flask* that allows users to track and visualize vehicle usage over time.
- Implemented a data model in a MySQL database using the star schema, supporting more than 1500 records.
- Incorporated a REST API comprised of 5 endpoints that allows users to authenticate and modify records.
- Designing an admin dashboard on Tableau, providing metrics such as the number of active users and monthly records.

### Citi Bike Inventory Tracker

Jan 2023 - May 2023

- Engineered an end-to-end application in Databricks to track Citi Bike inventory and weather in NYC.
- Built an immutable *PySpark* data pipeline to extract and refine data from various sources, including APIs.
- Forecasted hourly net bike change using *Prophet* and registered staging and production models using *MLflow*.

## Education

---

### University of Rochester

MS in Data Science, GPA: 4.00/4.00

Rochester, NY, USA

Jun 2022 - May 2023

### Clarkson University

BS in Mechanical Engineering, Minor in Mathematics, GPA: 3.84/4.00

Potsdam, NY, USA

Aug 2014 - May 2018

University of Leicester, study abroad (2 semesters)

Sep 2017 - May 2018

University of Technology Sydney, study abroad (1 semester)

Mar 2016 - Jun 2016