

# Jake Brehm

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

## Education

### University of Rochester

MS in Data Science

GPA: 4.00/4.00

Rochester, NY, USA

Jun 2022 - May 2023

**Relevant coursework:** Tools for Data Science, Database Systems, Computational Statistics, Data Mining, Data Science at Scale, Statistical Machine Learning, Network Science Analytics, Data Science Capstone & Practicum

### Clarkson University

BS in Mechanical Engineering, Minor in Mathematics

GPA: 3.84/4.00, with great distinction [🔗](#)

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

## Projects

### Specialty Referral Network [🔗](#)

Jan 2023 - May 2023

- Accentuated convolutions in UR Medicine's speciality referral network using *networkx* and *plotly*.
- Embedded the visualization into a *Flask* web app for additional customizability and layperson accessibility.
- Defined aesthetic parameters and custom efficiency metrics to determine where bottlenecks occur.

### NYC Citibike Inventory Tracker

Jan 2023 - May 2023

- Engineered a data-intensive application to track Citibike bike sharing inventory and weather in NYC.
- Built an immutable *PySpark* ETL 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*.

### European Tweet Classification [🔗](#)

Jan 2023 - Feb 2023

- Determined the ideologies of politicians with ~78% accuracy by analyzing multilingual tweets.
- Incorporated feature engineering techniques and a logistic regression model into an NLP pipeline.

## Experience

### Horizon Health Services

Data Engineer

Getzville, NY, USA

Jun 2023 - Present

- Strategizing with stakeholders and interdisciplinary team regularly using Agile principles.
- Building, testing, and maintaining ETL and data pipeline architectures using SQL and Python.
- Incorporating modern software practices including Git version control and scalable architectures.
- Spearheading development and maintenance of a data warehouse specifically designed for testing.

### Virufy

Associate Data Scientist

Remote • San Francisco, CA, USA

Jan 2023 - Present

- Collaborating with fellow volunteers to employ an AI that detects COVID from cough audio signals.
- Generating 100,000+ synthetic coughs using GANs to mitigate class imbalance in training data.
- Evaluating the diversity and quality of the synthetic data relative to real data using Inception Score.

### Viewpoint Systems Inc.

Systems Engineer

Buffalo/Rochester, NY, USA

Apr 2020 - Jun 2022

- Obtained 8 LabVIEW certifications covering advanced software architecture techniques.
- Conceived and supported a version of automated test software that runs without hardware.
- Delivered software that validates electrical functions of spacecraft thrust vector control systems.

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

Development Engineer, Development Engineering Intern

Liverpool, NY, USA

Aug 2018 - Jan 2020

- Automated control and testing of a digital motor controller, used by a team of 6.
- Launched software that allows for viewing and manipulating multiple sets of data simultaneously.
- Produced computer vision software that measures oil leakage with >90% accuracy using *OpenCV*.

## Skills

Python • R • SQL • SAS • Databricks • Spark • HTML • CSS • JavaScript • Git • SVN • Agile • Jira • AWS • ETL/ELT • BI  
• Excel • PowerPoint • Pandas • Matplotlib • Seaborn • Plotly • Tableau • scikit-learn • PyTorch • Flask