# Sam Kritchevsky

GitHub | Website | LinkedIn | sam.kritch@gmail.com | 336-414-8694 | NYC, NC, or Remote | Full-time or Contract

## Professional Summary

I am a software engineer and data scientist with 5 years of professional experience. I previously worked at SeatGeek, where I was the tech lead for a six-member Data Platform team after serving in various other roles across the data organization during my tenure. I am most at home in the domain of data engineering, but I also have plenty of experience with ML-eng, Data Science, Analytics, backend, and full-stack engineering—I often describe myself as a ”data generalist”.

Before my tech career I was enrolled in a Ph.D. program in physics, and I still maintain a hobbyist interest in physics, math, and math education. I occasionally write on these topics on my personal site.

Please note that I am currently re-entering the tech industry after a few years away. In the mean-time, I have been working on personal projects, writing, traveling, taking piano lessons, and working as a private math and physics tutor.

## Technical Skills

| | | |
|---|---|---|
| **Languages** | : | Python, SQL. Some Typescript, Javascript, Scala. |
| **Tech** | : | PostgreSQL, DBT, Spark, Flink, RabbitMQ, Elasticsearch, Redis, PostGIS, Pandas, Numpy, Tensorflow, Docker, Luigi, Dagster, Akka, React, Svelte, Docker, Nomad |
| **Cloud Tech** | : | Various AWS, incl. Redshift, Redshift Spectrum, Kinesis, Kinesis Firehose, S3, Lambda, EMR, Glue, IAM, RDS. |
| **Roles** | : | Data Engineering, Data Platform, Backend Engineering, ML Engineering, Data Science, Analytics |

## Experience

**SeatGeek** 2016-2021

*Sr. Software Engineer, Data Platform* 2020-21

Tech lead of the Data Platform team. Our projects included:

- A rewrite of our clickstream ingestion service, which massively improved scalability and fault tolerance (particularly during spiky event on-sales) and enabled fanout to many downstream applications. Tech: Flink, Spark
- Establishment of a standardized notebook stack, which allowed DS analysis work to be shared reproducibly across the team and to be incorporated into production data pipelines. Tech: Jupyter, Papermill
- Rollout of a centralized system for data lineage and discovery. Tech: Amundsen, Airflow, AWS Neptune.
- Ownership of the streaming data pipelines for ticket inventory, search, clickstream, and experimentation data. Tech: Spark, Akka Streaming, Python, AWS Kinesis.
- Ownership of the core analytics stack, backing our BI layer used by many departments. Tech: Druzhba, DBT, Luigi, Redshift, Looker.

*Senior Data Scientist* 2019-20

*Data Scientist* 2017-19

I specialized in user preference signals, our live event catalog, and push marketing. Projects included:

- Recommendation algorithms for weekly newsletter and various cart-abandonment and price-drop notifications, which tripled the revenue attributed to email notifications. Tech: Spark on AWS EMR, SQL.
- Event and performer popularity models, feeding into event recommendations and search. Tech: Tensorflow, Keras, Numpy.
- Rewrite of entity-linking algorithm, used to deduplicated users and devices for marketing attribution and funnel KPIs. Tech: SQL, Luigi.
- Ongoing involvement in the design and measurement of metrics for recommendations, search, and push marketing.

*Software Engineer, Discovery* 2016-17

- Backend engineering and data analysis for frontpage recommendations, search, social features, and push marketing.
- Built a reverse-ETL system to push datasets from our Data Warehouse into production PostgreSQL and MySQL databases.

**Other**

*Private Math Tutor, Club Z Tutoring* 2024-2025

*Physics Graduate Student and Teaching Assistant* 2014-2016

My research interests involved statistical mechanics and dynamical systems.

## Education

**University of Wisconsin** Madison, WI

*Physics Ph.D. Student* 2013-2015

**University of North Carolina** Chapel Hill, NC

*B.S. Physics, Math Minor* 2012