Sam Kritchevsky

Location: New York, NY

GitHub | LinkedIn | sam.kritch@gmail.com | 336-414-8694

ABOUT ME

I'm a data generalist and backend engineer, with 5 years of experience working across a modern cloud-based data stack, and a strong math background. I'm looking for a role where I can have impact with my current skillset, while gaining experience with full-stack data-intensive products, MLops, or statistical inference and forecasting.

TECHNICAL SKILLS

Languages	:	Python, Scala, SQL, some Javascript / Typescript
Technologies	:	PostgreSQL, Redshift, Elasticsearch, Redis, PostGIS, Spark, Flink, RabbitMQ, Akka, DBT, Docker, Nomad, Datadog, Grafana, AWS, EMR, Lambda, Kinesis, S3, Neptune, Jupyter, Luigi, Dagster
Roles	:	Data Science, Data Engineering, Data Platform, Backend, Analytics, ML Engineering

EXPERIENCE

Seatgeek   New York, NY
Sr. Software Engineer, Data Platform, Tech Lead   2020-2021

Tech lead of team of six engineers. Projects included:

- Rewrite of our in-house event ingestion service in Flink + Kafka to add features, modernize, and scale to bursts of high traffic. I advocated for the project, designed the system, and implemented much of the business logic.
  - ⋆ Fanouts to front page recommendations, push notifications, and Redshift, with plans to support an online feature store.
  - ⋆ Realtime sessionization, with handling of late-arriving data.
  - ⋆ Validation against a repository of JSONSchemas, with alerting in Datadog
- Deployed Amundsen, an OSS data-discovery platform, with a custom AWS Neptune backend.
- Jupyter infrastructure based on Docker, Papermill, and NBViewer.

We also owned the core data stack: Druzhba (built-in-house ETL), DBT, Looker, Luigi, Redshift, and other AWS products.

Sr. Data Scientist   2019-2020

Data Scientist II   2018

I specialized in user recommendation signals and our event catalog, working in Spark, SQL, and Python. Projects included:

- Algorithms for the weekly newsletter, cart abandonment and price drops notifications, in Spark + SQL. This tripled the direct revenue of the weekly newsletter immediately.
- Event and performer popularity models for event recommendations and search, built with Keras
- Entity-linking algorithm deduplicating users and device for marketing attribution and funnel KPIs.
- Ongoing involvement in the design and measurement of recommendation, search, and CRM marketing.

Software Engineer, Discovery   Sept 2016-2017

I built features for search, CRM marketing, and frontpage recommendations.

Other

Greater Harlem Coalition (Volunteer)   2022-
Contributed data analysis, writing, editing, and website administration for a community advocacy group. My data pipeline uses DBT, Dagster, and PostGIS (code here.)

Physics Teaching Assistant   2014-2016

EDUCATION

University of Wisconsin   Madison, WI
Coursework towards Physics Ph.D.   2013-2015

University of North Carolina   Chapel Hill, NC
B.S. Physics   2012