

Ryan Wallace

Boston, MA | ryan@ryancswallace.dev | 617-852-9239
github.com/ryancswallace | ryancswallace.dev | linkedin.com/in/ryancswallace

EDUCATION

Harvard University

Cambridge, MA

A.B. in Computer Science, cum laude; secondary concentration in Statistics; GPA: 3.8.

May 2018

Senior Research: “Modeling Behavior of Networked Agents with Dynamic Opinions”

Computing Coursework: Algorithms and Data Structures; Big Data Systems; Artificial Intelligence; Networks, Economics and Computation; Abstraction and Design in Computation.

Statistics Coursework: Real Analysis; Probability; Inferential Statistics; Time Series and Prediction; Generalized Linear Models; Econometrics.

EXPERIENCE

Federal Reserve Bank of Boston

Boston, MA

Lead Data Scientist, Research

January 2023 – February 2026

- Designed the architecture and coordinated implementation of an AWS data lakehouse supporting terabyte-scale research datasets and reproducible analytical workflows across the Federal Reserve System.
- Built Python/Spark tooling that accelerated terabyte-scale analysis, standardized researcher workflows, and automated administrative operations.
- Created and led an internal technical talk series to promote best practices in data science, machine learning, and software engineering.

Senior Data Scientist, Research

December 2021 – January 2023

- Architected and developed a data lake using Spark Streaming for ELT and Airflow for job orchestration.
- Categorized debit card transactions on multi-terabyte datasets using text clustering and classification techniques.

Data Scientist, Research

December 2020 – December 2021

- Partnered with economists to produce empirical reports on consumer spending response to the CARES Act. Developed a distributed lag regression model and robustness checks.
- Modeled expected effectiveness of the Main Street Lending Program (MSLP) using Fed warehouse data, SQL, scikit-learn, and Plotly.

DevOps Engineer, Research

December 2019 – December 2020

- Completed migration of compute, database, and web service infrastructure to AWS using IaC under CI/CD.
- Created data pipelines to continuously update datasets scraped from public sites and retrieved from REST APIs.

Harvard Data Ventures

Cambridge, MA

President

December 2017 – September 2018

- Provided technical and project management guidance on data science projects for Boston startups.
- Led training of fellows in weekly coding workshops, data science seminars, and an industry sponsorship program.

Fellow

September 2015 – December 2017

- Refined project specifications, developed models, completed analyses and simulations, and presented results regularly to clients' technical and business teams.

Arrowstreet Capital

Boston, MA

Research Intern

June 2017 – August 2017

- Implemented and deployed to production an equity risk analysis module based on a newly published model. Used Numba to achieve a 5-fold execution speedup over baseline.
- Applied Bayesian regression and time-series methods to decompose returns and support portfolio risk analysis.

BNY Mellon

New York, NY

Software Engineering Intern

June 2016 – August 2016

- Developed a tool to predict the likelihood of sales conversions using an ensemble model. Enabled BNY's sales team to sort and filter leads by estimated potential.
- Implemented a Salesforce client in Apex to consume and process REST API data from BNY's proprietary sales platform.

SKILLS

Languages: Python; SQL; Go; Bash; C; OCaml.

Data & ML Platforms: Spark; Airflow; AWS data lakehouse architecture; OLAP systems; stream processing; relational and key-value databases.

Modeling & Applied ML: Generalized linear models; Bayesian regression; time-series analysis; text clustering and classification; search and matching; model explanation; LLM application development.

ML Ops & Infrastructure: Model serving with Ollama and vLLM; Docker; Terraform; Ansible; Linux (RHEL, Alma); CI/CD; REST APIs; system administration automation; Git.

Observability: Prometheus; Grafana; CloudWatch.

Engineering Practices: Object-oriented and functional programming; agile development; production data tooling; technical project management.