Christopher Donnay

christopherdonnay@gmail.com | cdonnay.github.io

Summary

Mathematics PhD with five years experience specializing in data science and statistical modeling. Demonstrated expertise in solving high-impact problems with technical solutions and designing, implementing, and validating statistical models in Python. Strong ability to manage projects, communicate high-level mathematics to non-technical audiences, and provide actionable insights. Seeking a position as a data scientist or analyst.

Skills & Certifications

- **Mathematics**: Data science, machine learning, probability, statistics, statistical inference, topological data analysis, Markov chains, graphs and networks, topology, combinatorics, discrete geometry, gerrymandering, voting methods.
- **Languages & Platforms**: Python, SQL, C++, Docker, AWS (S3, Lambda), HPC & Slurm, Bash, GitHub, QGIS, Mathematica, LaTeX, Svelte(kit), Sphinx, Markdown.
- **Python libraries**: Jupyter, Matplotlib, NetworkX, NumPy, (Geo)Pandas, Poetry, seaborn, scikit-learn, uv.
- **Certifications**: The Erdős Institute Data Science Boot Camp.
- **Soft skills**: Leadership, self-motivation, project management, (science) communication, problem solving, open source collaboration, attention to detail, teaching.

Experience

Data and Democracy Lab — Remote

Lab Manager — January 2025 - Present

- Lead a team of six contributors to `VoteKit` (votekit.readthedocs.io), the Lab's Python package for statistical modeling and election analytics. Design, prototype, evaluate, and implement three new statistical models of voter behavior.
- Product manage the development and public release of Districtr 2.0 (beta.districtr.org), a browser-based geospatial app for drawing legislative districts, with a remote team of five full-stack developers. Used by over 150 localities and states as their official mapping tool for redistricting.

Research Scientist — January 2024 - January 2025

- Created a model of redistricting plans in Michigan, processing large-scale geospatial and electoral data in order to infer the impact of the Voting Rights Act on partisanship. Results were presented to the advocacy group Voters Not Politicians to influence creation of new legislative maps for the Michigan Supreme Court case Agee v. Benson.
- Designed, wrote, and executed a program in statistical inference of polarized voting run by the Southern Poverty Law Center. Developed six Jupyter notebooks in Python with data processing and modeling exercises.

Selected Projects

- **"3:1 Nesting Rules in Redistricting"**: Using Markov chain methods, coded a novel statistical model for nested legislative maps in Ohio and Wisconsin with Python. Model was used to detect anomalous gerrymanders and predict and recognize patterns in the electoral outcome of maps. Validated the model using mixing heuristics and other tools. Invited to present at SIAM Annual Meeting, July 2024, and is set to appear in the journal *Statistics and Public Policy*.
- **"Portland, OR 2024 Election Analysis"**: Provided Python support to an analysis of the recent STV election, including processing of ballots with `pandas` and developing visualizations with `matplotlib`. Showed that dominant media narratives about ballot errors by people of color were misguided. Presented findings five times throughout the spring of 2025 to community groups, academics, and non-profits. White paper available at mggg.org/ppm.
- **"Predicting the Winner of the Great British Baking Show"**: Trained supervised learning models using `scikit-learn`—regression, random forest, k-nearest neighbors, and Naive Bayes—to predict winners and uncover key drivers of success in GBBS. Presented to a panel of industry experts at the Erdős Institute Data Science Bootcamp, who specifically highlighted our team's clear communication of modeling limitations and results. Awarded first place.

Education

The Ohio State University — PhD Mathematics, 2024

The Ohio State University — MS Mathematics, 2024

University of Pennsylvania — MS Education, 2020

Pomona College — BA Mathematics, 2018