

Kalynn H. Willis

2874 Gateway Avenue · Hartford, WI 53027 · 262-224-8625 · kalynnhopewillis@gmail.com

EDUCATION

August 2026

University of Wisconsin-Madison, Madison, WI

Intended Degree: Bachelor of Science

Intended Majors: Statistics & Data Science

GPA: 3.92

Honors & Awards: Dean's List (5 semesters)

Relevant Projects: Performed statistical analysis of CDC COVID-19 surveillance data (April 2021–Sept 2022, 72% U.S. population coverage) to evaluate Moderna vs. Pfizer vaccine death rates; applied z-tests for proportions, confidence intervals, and incidence-rate trend analysis in R and Python to demonstrate significant differences in mortality risk.

EXPERIENCE

AprilAire

January 2026 - Present

Data Engineer Co-op

Madison, WI

- Building Merv, a customer-facing **AI Assistant** that delivers personalized healthy-air product recommendations by combining a conversational LLM interface with an **XGBoost Model** that predicts optimal product, model, and wattage configurations based on home characteristics; **currently in development for integration into the AprilAire website and mobile app**
- Designed and deployed an **MCP server** that automates firmware-level decisions for connected thermostats and air-quality devices, dynamically adjusting setpoints in response to real-time conditions and forecasted weather events (e.g., extreme overnight cold)
- Built a company-wide "heartbeat" monitoring system that tracks overnight job execution and surfaces failures to engineering teams, improving observability and reducing time-to-detection for pipeline issues
- Engineered ETL pipelines in **AWS** (python) to ingest, transform, and warehouse high-volume IoT device telemetry, supporting analytics and ML workflows across the company
- Conducted statistical analysis to evaluate the energy efficiency of Healthy Air mode in vent systems, informing product decisions with data-backed performance comparisons

University of Wisconsin-Madison Department of Biochemistry.

October 2024 - December 2025

Data Science Undergraduate Researcher

Madison, WI

- Led development of a **RAG-style research chatbot** leveraging **FastAPI**, **vector embeddings**, and **local/remote LLMs**, allowing researchers to query phenotype data in natural language; **supervised** an undergraduate student contributing to development and testing
- Implemented **API** integrations with public genetic resources (Ensembl, JAX, GTEEx, IMPC), ensuring accurate, up-to-date annotations across multiple species
- Designed and maintained interactive visualization platforms with **RShiny**, **Next.js**, and **TypeScript**, supporting real-time exploration of QTL peaks, gene annotations, and phenotype trends
- Developed and deployed containerized applications with **Docker** and **HTCondor**, enabling reproducible, distributed workflows for high-throughput genetic studies
- Engineered large-scale data pipelines for quantitative trait locus (QTL) and gene mapping analyses, integrating DuckDB and SQLite databases to query billions of genetic records efficiently
- Collaborated with biologists, statisticians, and computational researchers to maintain data integrity, enforce version control (GitHub), and translate complex results into usable tools for lab-wide and future public adoption

Arrowhead Pharmaceuticals

May 2025-August 2025

Data Science Intern

Madison, WI

- Assisted the AI/ML team in development of machine learning models (**gradient boosting**, **gradient descent-based methods**) to analyze and predict molecular and genetic properties relevant to therapeutic discovery
- Collaborated with computational scientists to evaluate model performance, optimize hyper-parameters, and compare alternative ML approaches for biological datasets
- Applied **Python**, **Pandas**, and **scikit-learn** to preprocess large datasets, streamline feature engineering, and validate predictive results
- Developed an automated **ETL** pipeline in Python with **CI/CD** integration to load large experiment datasets into a centralized research database, reducing manual data prep time by 50% and ensuring reproducibility using **Github**.
- Contributed to ongoing research projects bridging bioinformatics and drug discovery, enhancing the team's ability to generate insights from high-dimensional molecular data

Volunteering

Actuarial Science Club

September 2024

Volunteer Committee Member

Madison, WI

- Handpicked to lead, coordinate and organize volunteer and community engagement events
- Worked closely with committee members and external partners and clubs to successfully execute multiple volunteer initiatives

Madison Metropolitan School District: West High School

October 2023-May 2024

Achievement Connections Advanced Math Tutor

Madison, WI

- Devoted 8 hours per week to organize and lead resource period sessions, collaborating closely with teachers to ensure effective coordination of learning objectives and facilitating large group tutoring sessions

- Provided comprehensive support in advanced math subjects to high school students, with a focus on clarifying complex calculus topics

Junior State of America

October 2020 - June 2022

Founder & President

Slinger, WI

- Led bi-weekly meetings and coordinated debate topics, ensuring a safe space for peers to have productive conversations
- Established a leadership Cabinet and delegated responsibilities to members, fostering teamwork and empowering peers to take on leadership roles

SKILLS

Technologies: Python, R, Typescript, AWS, SQL, FastAPI, Ollama, OpenAI, Docker, Rshiny, Next.js, PyTorch, Condor, Cursor, TailwindCSS

Skills: Server management, Team Leadership, AI agent architecture, end-to-end development

Interests: Reading; Golf; Traveling; Running