

Matthew Kim

Fairfax, VA | (571) 351-7450 | mjuwon11@gmail.com
linkedin.com/in/matthewkim1104 | github.com/mattkoom11 | juwonkimportfolio.org

EDUCATION

University of Virginia

B.A. in Applied Statistics, Minor in Computer Science | GPA: 3.9/4.0

Charlottesville, VA

Aug. 2023 – May 2027

- **Relevant Coursework:** Statistical Machine Learning, Data Visual and Management, Data Analysis with Python, Design & Analysis of Sample Surveys, Intro Mathematical Statistics, From Data to Knowledge, Intro to Regression Analysis, Software Dev Essentials, Computer Systems and Org 1, Discrete Math and Theory 1, Data Structures and Algo 1

EXPERIENCE

Research Assistant

UVA Historic Fashion Archive

2024 – Present

Charlottesville, VA

- Architecting a Next.js 16 / React 19 / TypeScript frontend for UVA's vintage clothing collection, integrating a CollectiveAccess CMS through a custom API client with cookie and Basic Auth token flow, a 5-minute in-memory cache, and a static JSON fallback to ensure data availability under varying backend conditions
- Implementing a content-based machine learning algorithm that scores garment similarity across five weighted attributes, including era, type, color, material, and decade, to power related-item recommendations without a dedicated ML backend
- Building the full collection data pipeline with a paginated grid, multi-attribute filtering, advanced Boolean search, analytics tracking, and JSON/CSV/PDF export; designing an interactive 3D garment viewer using Three.js / @react-three/fiber
- Piloting a statistical pipeline that extends the recommendation algorithm with geometry-based features by extracting shape descriptors, including curvature, symmetry, and drape, from 3D garment meshes via `trimesh`, then applying PCA and clustering in Python and R with model evaluation against catalog metadata using ARI and NMI

Teaching Assistant

University of Virginia — CS1112

2026 – Present

Charlottesville, VA

- Supporting 300+ students in introductory Python through weekly office hours, lab sections, and one-on-one debugging
- Grading assignments and exams with structured feedback on data structures, control flow, and functions

PROJECTS

Hana-Bi: E-Commerce Platform | Next.js, TypeScript, Stripe, Resend, Zustand, Framer Motion

2024 – Present

- Engineering a full-stack e-commerce platform with Stripe as the product data source of truth, implementing an autopaging catalog pipeline, a per-product metadata schema, and serverless-safe session verification via direct Stripe API calls
- Building an end-to-end transactional data flow: client cart state (Zustand, persisted locally), Checkout Session creation, webhook-driven order confirmation emails via Resend, and a waitlist capture pipeline with dual-notification routing
- Structuring a multi-environment deployment checklist covering API key rotation, webhook secrets, and sender domain verification

Tech Industry Layoff Analysis | R, tidyverse, ggplot2, R Markdown

2026

- Analyzed 2,700+ tech layoffs using R and tidyverse, applying statistical modeling techniques including t-tests and ANOVA to compare workforce reduction severity across company types and industries
- Built exploratory visualizations in ggplot2 and engineered features from raw data; evaluated workforce reduction patterns and found that private companies cut proportionally twice as much of their workforce as public ones (32.4% vs. 16.7%)
- Collaborated in a 3-person team using GitHub for version control, delivering a fully reproducible R Markdown report

Insurance Client Data Management System | Python, Flask, SQLite, Google Sheets API, Render

2025 – 2026

- Built a Flask dashboard that syncs the agency's Google Sheet into a read-only SQLite copy for search, sorting, and CSV export, replacing a binder-based workflow covering ~2,000 records; rebuilt in 2026 with a 30/60/90-day renewals view, an audit log, and a pytest suite in GitHub Actions

NBA Salary Prediction Model | R, multiple linear regression, cross-validation

2025

- Built a multiple linear regression model predicting NBA player salaries from performance, awards, age, and position; performed full EDA, variable selection, and assumption diagnostics (linearity, homoscedasticity, normality of residuals)
- Achieved adj. $R^2 \approx 0.77$ via five-fold cross-validation; identified Points, Minutes, Rebounds, and Assists as dominant predictors; flagged high-leverage observations via Cook's distance and residual analysis

TECHNICAL SKILLS

Languages: Python, R, Java, C, JavaScript, TypeScript, SQL

Frameworks & Libraries: React, Next.js, Flask, Tailwind CSS, pandas, NumPy, tidyverse, ggplot2, Three.js, Generative AI, Retrieval-Augmented Generation (RAG)

Developer Tools: Git/GitHub, Stripe API, Google Sheets API, Resend, Zustand, Render, Jupyter, REST APIs, MLOps/LLMOps, CI/CD

Statistical Methods: Hypothesis testing, ANOVA, regression, bootstrap, Monte Carlo simulation, cross-validation, EDA, Machine Learning, Optimization, Forecasting, Model Evaluation