# Luke Spencer

(925) 523-9466 | lukespencer2006@gmail.com | linkedin.com/in/lukehspencer/ | github.com/lukehspencer

## EDUCATION

**University of California, Berkeley** **Berkeley, CA**
*Bachelor of Arts, BA - Applied Mathematics & Data Science* *GPA: 3.87*
- **Relevant Coursework:** Structure and Interpretation of Computer Programs (CS 61A), Linear Algebra and Differential Equations (Math 54), Foundations of Data Science (Data C8), Multivariable Calculus (Math 53)

## WORK EXPERIENCE

**UC Berkeley, Rausser College of Natural Resources** **Berkeley, CA**
*Data Science Researcher* *Sep. 2025 – Present*
- Process and clean 5M+ rows of Colombian migration, individual, and household data from 2012-2021 using Python and R/RStudio to ensure accuracy for downstream analysis and consistency across datasets
- Constructed reproducible data pipelines to standardize data structures, and derived relevant variables such as migration flows, variable weights, and individual characteristics for climate-migration analysis

**iCode School** **Pleasanton, CA**
*Coding Instructor* *Apr. 2025 - Aug. 2025*
- Taught programming and gaming to students aged 5-10 using structured curricula and hands-on projects
- Enhanced engagement and retention of course material by adapting lessons and teachings to learning styles
- Increased program enrollment 15% by hosting STEM workshops and marketing initiatives for students

**UC Santa Barbara, Bren School of Environmental Science & Management** **Santa Barbara, CA**
*Data Science Researcher* *Jun. 2024 - Jul. 2024*
- Built multi-species Lotka–Volterra ecological models in R/RStudio, forecasting 50 years of species growth to evaluate ecosystem stability and derive insights into parrotfish grazing effects on algae and coral reefs
- Co-authored a 10 page research paper and delivered an academic presentation synthesizing modeling assumptions, simulation results, and long-term population dynamics into clear ecological results

## PROJECTS

**Elytra** **Oct. 2025**
*Python, TypeScript, Next.js, HTML/CSS, FastAPI, Pandas, Vercel, Render*
- Architected deployment pipeline for full-stack AI platform that identifies the most effective AI tools for a company's problem by analyzing 13K+ AI tools; built tool reports using LLM and heuristic calculations
- Assisted in API integration between Next.js/TypeScript frontend and FastAPI/Python backend, resolving production CORS configuration and multi-platform deployment to enable intelligent tool matching

**Harvard CS50AI** **July. 2025 – Aug. 2025**
*Python, Scikit-Learn, TensorFlow*
- Built and trained a convolutional neural network with TensorFlow to classify 43 traffic sign classes across 23K+ images, achieving 98% accuracy through preprocessing and systematic hyperparameter tuning
- Implemented a reinforcement learning agent using Q-learning that achieved >95% win rate in the game of Nim after 50K+ self-play iterations, converging to optimal strategies via iterative state–action value updates

**NBA MVP Predictor** **Jun. 2025**
*Python, Selenium, Beautiful Soup, Pandas, Scikit-Learn, Matplotlib*
- Developed a reproducible web scraping pipeline for Basketball Reference data using requests and Selenium, parsed HTML with Beautiful Soup, and prepared 15K+ player, team, and MVP data from 1991 to 2024
- Constructed a Random Forest model with 12+ engineered features (per-game stats, team performance data, etc.), achieving a mean average precision of 75.5% using an expanding-window backtest from 1996 to 2024

## SKILLS & INTERESTS

- **Languages:** Python, SQL, R, Java, HTML/CSS, TypeScript
- **Frameworks & Libraries:** Flask, FastAPI, Next.js, TensorFlow, Pandas, NumPy, Matplotlib, Scikit-Learn
- **Developer Tools:** Git, Jupyter Notebook, VS Code, RStudio, Hugging Face, Vercel, Render, Supabase
- **Interests:** Soccer, Football, Comedies, LEGOs, Hiking, Fitness, Travel, History, Geography, Star Wars