

# Ju Ho Kim

(352) 709-3825 | [juho\\_kim@berkeley.edu](mailto:juho_kim@berkeley.edu) | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

## SKILLS

**Data Science & Machine Learning:** Predictive Modeling, Statistical Analysis, Machine Learning (Regression, Classification, Clustering), Natural Language Processing (NLP), Model Evaluation  
**Data Preparation, Analysis, and Reporting:** Exploratory Data Analysis (EDA), Feature Engineering, Data Cleaning, Data Wrangling, Visualization  
**Data Engineering:** ETL Pipelines, Schema Design, Data Lineage, Web Scraping  
**Programming & Databases:** Python (Pandas, NumPy, Scikit-learn, Streamlit, Joblib), SQL (PostgreSQL)  
**Version Control & Collaboration:** Git, GitHub

## EXPERIENCE

- Data Science Undergraduate Researcher** Jan 2026 – Present  
*Lawrence Berkeley National Laboratory* *Berkeley, CA*
- Architecting a unified SQL database and Dash dashboard for ENIGMA (10-year national research project) to centralize disparate datasets, enabling multi-institutional scientific teams to streamline data discovery and exploration
- Data Analyst Intern** Jun 2024 – Aug 2024  
*Global ESG Foundation* *New York, NY*
- Built and enhanced a decision-ready Streamlit dashboard by wrangling and transforming global CO<sub>2</sub> emissions data into stakeholder-oriented metrics (country, absolute volume, growth rates), enabling clear, actionable insights for ESG analysis and decision-making
  - Successfully presented the solution to leadership in a quarterly board meeting and delivered the tool as a maintainable Git repository, enabling continued use and iteration beyond the internship
- Medical Evacuation Coordinator** Nov 2020 – Apr 2022  
*Sixth Division, Republic of Korea Army* *Gangwon, South Korea*
- Provided rapid-response medical support for demining operations in high-risk border zones, coordinating with UN forces under GPS- and communication-denied conditions to ensure mission-critical evacuations
  - Identified a division-wide gap in navigation resources and independently created a route-mapping system for drivers; adopted across multiple units, earning a commendation from a two-star general

## PROJECTS

- News Article Political Bias Classifier** | *Python, PyTorch, Hugging Face, NLP, Web Scraping*
- Achieved 69.7% classification accuracy on 24K+ news articles spanning 13 years, earning top-project recognition and a graduate TA-level, production-ready evaluation
  - Built and scaled an end-to-end NLP pipeline from raw HTML to model-ready text, designing robust preprocessing workflows (web scraping, normalization, deduplication) to ensure reproducibility and model reliability
- Consumer Default Risk Modeling** | *Python, Scikit-learn, Logistic Regression, Decision Trees*
- Predicted financial default risk using a complete ML modeling workflow (Logistic Regression, Decision Trees), earning strong faculty recognition for modeling decisions and evaluation
  - Enhanced model interpretability by identifying 'Balance' as the key risk driver through Sigmoid curve visualization and streamlined deployment using Joblib for model persistence
- Global CO<sub>2</sub> Emissions Dashboard** | *Python, Streamlit, Plotly, Pandas, Data Visualization*
- Developed an interactive dashboard to explore global CO<sub>2</sub> emissions trends (1960–2023) with volume and growth analysis across 205 countries
  - Implemented multi-layer data cleaning to exclude non-country aggregates, enabling accurate country-level comparisons

## AWARDS

- Honors in Mathematics Scholarship Award (Dean's List)** | *Pasadena City College* Jun 2025  
**Commendation for Operational Excellence** | *Sixth Division, Republic of Korea Army* May 2022

## EDUCATION

- University of California, Berkeley** Expected May 2027  
*B.A. Statistics, College of Computing, Data Science, and Society* *Berkeley, CA*
- Pasadena City College** Aug 2023 – May 2025  
*Associate of Science, Computer Information Systems* *Pasadena, CA*