

# Yeonji Baek

ybaek@andrew.cmu.edu | (714) 356-8700 | linkedin.com/in/yeonji-baek/

## EDUCATION

### Carnegie Mellon University

Bachelor of Science in Statistics and Machine Learning; Minor in Business Analytics

Pittsburgh, Pennsylvania

August 2022 - May 2026

**Cumulative GPA:** 3.52

**Relevant Coursework:** Principles of Imperative Computation, Algorithms and Advanced Data Structures, Probability and Statistical Inference, Applied Machine Learning, Applied Data Management, Optimization for Business Analytics, Modern Regression

## WORK EXPERIENCE

### Horizon Advisory

Pittsburgh, Pennsylvania

Data Engineering Intern

June 2024 – August 2024

- Analyzed dataset on China's potential academic leaders, inventing new categories to resolve company's data quality issues
- Automated information extraction for new categories, demonstrating improved dataset utility to support suggested enhancements
- Enhanced web scraper tool by implementing parsing algorithms to extract data from unique websites to improve data quality
- Presented comparative analyses showing advantages of web-scraped dataset over the original, emphasizing its future potential

### CMU Concepts, Actions, and Object Lab (CAOs)

Pittsburgh, Pennsylvania

Research Assistant

August 2023 – May 2024

- Developed a program in Python utilizing OS, NumPy, and OpenCV to extract thousands of frames from videos for a dataset
- Implemented neural network models using Fastai2 library, evaluating various model accuracy using original and altered frames
- Designed research program on PsychoPy to test intuitive physics in participants utilizing the same frames tested with the models
- Facilitated experiments on 40+ participants to collect data to compare the difference between human and models' accuracy

## PROJECTS

### Personal Project: LookUp Web Application Development

November 2024 – Present

Developer | Figma, Valentina Studio, HTML, CSS, JavaScript

- Designed interactive prototypes and database schema, establishing a blueprint for application's structure and user experience
- Developing a client-side interface with HTML, CSS, and JS, focusing on responsive design and seamless user interactions

### Asian Student Association Website Redesign

June 2024 – Present

Webmaster | <http://www.cmuasa.org/> | HTML, CSS, JavaScript

- Redesigning the website to appeal to first-years during orientation to solve the problem of decreasing number of members
- Transformed the structure from a single scrolling page to a multi-page format for better content organization and navigation

### “Traffic Stops in Hartford, Connecticut” Project

November 2023 – December 2023

Statistical Graphics & Visualization Group Project | <https://tinyurl.com/39tsmvfc> | R

- Conducted in-depth analysis on a dataset of 18k+ using R (Tidyverse, ggplot), focusing on geographic distribution of traffic stops
- Generated interactive graphs, identifying a potential trend of higher frequency of arrests that can be generalized to other areas

## LEADERSHIP EXPERIENCE

### Asian Students Association

Pittsburgh, Pennsylvania

Vice President of Finance

October 2023 – Present

- Optimized fundraising strategies through data analysis, identifying high-profit items and setting a revenue record of \$3000+
- Managed a budget of \$50k to plan events for 600+ members, coordinating ASA Night Market featuring 16 cultural organizations

### Carnegie Mellon University First-Year Orientation

Pittsburgh, Pennsylvania

Head Orientation Counselor

February 2024 – Present

- Oversaw 130 Orientation Counselors, equipping OCs to guide 1000+ incoming students through transition into university life
- Redesigned orientation program structure by executing engagement strategies, boosting in 23% increase from last year

### Data Science Club

Pittsburgh, Pennsylvania

Social Engagement Director

September 2024 – Present

- Introduced RSVP calendar system to track engagement, investigating patterns and increasing overall participation by 17%
- Established key partnerships with university departments and external organizations, hosting datathon with 200+ participants

## SKILLS / INTERESTS

**Programming Languages:** Python (NumPy, Pandas, Scikit-Learn), C++, R, SQL, HTML, CSS, LaTeX

**Technologies:** Tableau, MongoDB, Microsoft Excel, Adobe Creative Suite (Photoshop, InDesign, Lightroom), Git/GitHub

**Languages:** English (Advanced), Korean (Advanced)

**Interests:** Sports Photography, Volleyball, Learning Foreign Languages (Japanese, Chinese), Sudoku