

# ETHAN LEE

781-690-7469 • ethan\_lee@college.harvard.edu • github.com/eyechan01

## EDUCATION

### Harvard University

Cambridge, MA

A.M in Statistics

May 2023 (expected)

A.B in Computer Science and Statistics, GPA: 3.885/4.00

May 2023 (expected)

**Relevant Coursework:** Quantitative Finance. Statistical Inference. Probability. Machine Learning. Data Structures & Algorithms.

Systems Programming & Machine Organization. Quantum Computation & Quantum Complexity.

**Activities:** Harvard Open Data Project (President). Harvard Computer Society (Director of Diversity & Inclusion).

Datamatch (Algorithms/Web Dev Specialist). Harvard Crimson. Satire V. Asian American Dance Troupe.

### Needham High School

Needham, MA

US Presidential Scholars Semifinalist. Superintendent's Academic Excellence. Class Day Speaker. AIME Qualifier.

June 2019

## WORK EXPERIENCE

### Capital One

New York, NY

Intern, Financial Software

June 2022 – Present

- Developing unified testing environment in Python and JavaScript for simulating trading algorithms, regression tests, and other unit tests to be used in Capital One's Capital Markets division.

### World Data Lab

Vienna, Austria

Principal Investigator, Econometric Modeling and Policy Research

August 2021 – March 2022

- Created econometric modeling algorithm for classifying economic and financial factors in NYC data; to be used to affect policies on financial recovery from COVID-19 in the Bronx.
- Selected by Alfred P. Sloan Foundation for [research grant](#) through "extremely rigorous" review.

### Republic of Paraguay, Department of Public Contracting

Asuncion, Paraguay

Intern, Financial Modeling

June – September 2021

- [Predicted](#) market prices of goods/services in Paraguay's public procurement process using R and Python.
- Used linear regression, random forest, XGBoost, generalized additive model, and elastic net regression; reduced RMSE to within 3 USD of actual prices.
- Deployed model through RShiny to be implemented in Paraguay financial corruption investigations.

### Echelon Insights ([echeloninsights.com](https://echeloninsights.com))

Washington, D.C.

Intern, Data Science

June – August 2020

- Sentiment Analysis Project: Created sentiment analysis models for business discourse using natural language processing and neural networks with Python; improved model accuracy from 15% to 60%.
- Political Data Scraping Project: Scraped election result data from 100+ congressional districts through BeautifulSoup.

### MIT Computer Science & Artificial Intelligence Lab

Cambridge, MA

Intern, Functional Programming

June – July 2019

- Proved simulation property for [Hemiola](#), a cache-coherence protocol design and proof framework, using Coq.

## LEADERSHIP & ACTIVITIES

### Trends in Cryptocurrency Prices

Cambridge, MA

Project Leader

July 2022 – Present

- Using Python libraries and deep learning models to analyze public cryptocurrency data and discover factors influencing cryptocurrency prices as well as ways to forecast their future fluctuations.

### Harvard Investment Portfolio Analysis

Cambridge, MA

Project Leader

May – July 2022

- Used Python to conduct time series [analysis](#) and evaluate beta, delta, and other aspects of Harvard investment portfolio over 20 years to understand its performance and determine its most contributing factors.

### Harvard Open Data Project

Cambridge, MA

President

March 2021 – April 2022

- Led and managed 250+ members and \$15k+ in funds.
- With Cambridge Open Data Program, organized annual datathon; led [case competition](#) with 150+ participants on analyzing housing market and affordability in Cambridge; reports presented to Cambridge zoning and housing staff.
- Oversaw renovation and update of [hodp.org](https://hodp.org), hosted in Sanity and Gatsby.

### Philippines Job Market Analysis

Cambridge, MA

Project Manager

November – December 2020

- Led [analysis](#) of factors of first-time salaries in the Philippines, investigating effects of industry, gender, and university.
- Used linear regression, mixed effects, LASSO, and random forest models; RMSE of best model as low as 0.18.

## SKILLS & INTERESTS

**Technical Skills:** Python, R, SQL, JavaScript (React, Node/Express, D3), C++, Java, HTML/CSS, MATLAB, OCaml, and Coq.

**Personal Interests:** Applied Machine Learning, Financial Analysis, Soccer Transfer Market, Story Writing, Music Composition.