# ETHAN LEE

781-690-7469 • ethan_lee@college.harvard.edu • github.com/eyechan01

## EDUCATION

**Harvard University**                                                                                     Cambridge, MA
A.B in Computer Science and Statistics, GPA: 3.88/4.00                                        May 2023 (expected)
**Relevant Coursework**: Intro to Statistical Inference. Linear Models. Probability. Data Structures and Algorithms. Abstraction & Design in Computation. Systems Programming & Machine Organization.
**Activities:** Harvard Open Data Project (President). Harvard Computer Society (Director of Diversity & Inclusion). Datamatch (Algorithms/Web Dev Specialist). Harvard Crimson. Satire V. Harvard Breakers.

**Needham High School**                                                                                   Needham, MA
US Presidential Scholars Semifinalist. Superintendent's Academic Excellence. Class Day Speaker. AIME Qualifier.        June 2019

## WORK EXPERIENCE

**World Data Lab**                                                                                          Vienna, Austria
**Principal Investigator, Econometric Modeling and Policy Research**                  August 2021 - Present
- Developing 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 2021 - Present
- Predicting 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.

**Harvard University, Department of Psychiatry**                                        Cambridge, MA
**Research Assistant, Decision-Making & Mental Disorder Research**                  June 2021 – Present
- Using Python machine learning methods and network analysis on neuroimaging datasets with McLean Hospital Applied Neuroimaging Statistics Laboratory to study decision-making and mental disorders.

**Echelon Insights** ([echeloninsights.com](http://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

**Harvard Open Data Project**                                                                       Cambridge, MA
**President**                                                                                           March 2021 – Present
- Leading and managing 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](http://hodp.org), hosted in Sanity and Gatsby.

**VotrCrwd, IvyHacks Competition** ([votrcrwd.netlify.app](http://votrcrwd.netlify.app))                              Cambridge, MA
**Project Leader**                                                                                   October 2020 – May 2021
- Implemented location API and frontend for voting location capacity and information website in React and Firebase.
- Won IvyHacks "Best Use of Google Cloud – COVID-19 Hackathon"; expanded at Google Cloud COVID-19 Hackathon.

**Global Health Budget & Policy Visualization**                                        Cambridge, MA
**Project Leader**                                                                                   October– December 2020
- Created interactive [webpage](#) on global health budgets for disease prevention as proportions of GDP expenditures.

**Philippines Job Market Analysis**                                                                Cambridge, MA
**Project Manager**                                                                                 November – December 2020
- Led [analysis](#) of factors of first-time salaries in the Philippines, investigating factors such as 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, JavaScript (React, D3), C++, Java, HTML/CSS, MATLAB, SQL, OCaml, and Coq.

**Personal Interests:** Soccer Transfer Market, Short Story Writing, Tetris, Guitar, Piano.