

Yihan Shi

(919) 884-0734 | yihan.shiwork@gmail.com | yihan-shi.github.io

Education

Duke University

Durham, NC

Bachelor of Science (*cum laude*); Major in Statistics with High Distinction

Aug 2019 – May 2023

- Honors Thesis: *Investigate the impact of Confounding Medical Interventions (CMI) in electronic health records (EHR) data*

Experience

The Brookings Institution, Center on Health Policy

Washington, DC

Senior Research Assistant

Aug 2024 – Present

- Co-authored research papers and comment letters with economists and lawyers on the Inflation Reduction Act, pharmaceutical import security investigations, and health insurance price transparency
- Built a data pipeline integrating historical FDA drug data and developed a dashboard visualizing U.S. drug shortage trends
- Led the full intern hiring process and provided mentorship and supervision to the 2 selected interns

Mastercard Data & Services

Arlington, VA

Software Engineer

Aug 2023 – Aug 2024

- Led an innovative research project using artificial neural networks to predict counterfactual campaign outcomes
- Built full-stack web application features to serve enterprise customers such as McDonald's, Walmart, and BoA
- Partnered with product managers to revamp the application frontend to meet industry-leading accessibility standards

Duke University, Department of Statistical Science

Durham, NC

Undergraduate Research Assistant

Aug 2021 – May 2023

- Co-authored an internal tutorial on recurrent event analysis examining alcohol-related complications after bariatric surgeries
- Investigated algorithmic bias of inappropriate proxy variables in Electronic Health Record data using simulation

Undergraduate Teaching Assistant

Aug 2022 – Dec 2022

- Held weekly office hours to help more than 70 students understand statistics concepts and answer coding questions
- Guided 2 teams of students in the design of a data analysis project, which involved cleaning online datasets, generating hypotheses, and evaluating statistical models

Carta Healthcare

San Francisco, CA

Data Science Intern

May 2022 – Aug 2022

- Built a client-facing dashboard tracking hospital performance metrics and presented it in a live product demonstration
- Created a reusable template that enabled engineers to auto-generate webpages for a machine learning workflow
- Implemented an automated data-cleaning pipeline that reduced bioinformaticians' housekeeping time by 30%

Duke University School of Medicine

Durham, NC

Summer Research Fellow

May 2021 – Aug 2021

- Co-led a team of 3 undergraduate and postdoc students to develop a rare disease genomic data science project
- Identified 20+ significant genes associated with a rare neurological disorder from a pool of 35,000 using statistical methods
- Mentored 2 incoming first-year researchers and designed introductory biomedical data science projects

Pea Rare Disease Initiative

Shanghai, China

Volunteer

Jul 2020 – Jan 2021

- Analyzed FDA orphan drug approval data to inform patient families about trends in U.S. orphan drug development
- Produced quarterly social media newsletters to raise awareness of rare diseases in China

Skills and Interests

- **Languages:** Mandarin Chinese (native), Japanese (intermediate)

- **Technical Skills:** Python, R, SQL, C#, C, Java, Assembly, JavaScript, CSS, HTML, L^AT_EX

- **Interests:** Photography, Piano, Pilates