Hello, I'm

# Yingqi Gao

8586667501
yingqi.gao.stats@gmail.com
linkedin.com/in/yingqi-gao
yingqi-gao.github.io

PhD candidate in statistics researching machine learning and economics, building real-world, data-driven solutions that help corporations maximize profits by transforming unstructured data and complex insights into strategic decisions. I develop market-driven mechanisms for digital economies, spanning data, privacy, and AI derivatives; and design AI systems that adapt to economic incentives, including LLM-powered ad auctions and uncertainty-aware chatbots.

## Education

**University of California, Los Angeles** - *PhD in Statistics & Data Science* SEP 2021 - JUN 2026

**Columbia University in the City of New York** - *MA in Statistics in Statistics* SEP 2019 - DEC 2020

**University of California, San Diego** - *BS in Probability & Statistics and BS in Management Science* SEP 2015 - JUN 2019

## Experience

**PhD Researcher** - *UC Los Angeles* FEB 2023 - PRESENT

- Pioneered a two-phase pricing mechanism that maximizes customer retention while ensuring scalable and revenue-optimal dataset sales in dynamic markets with fluctuating buyer participation.
- Engineered a first-of-its-kind online learning algorithm that accelerates adaptive pricing and regret reduction by leveraging insights from similar datasets, further driving long-term revenue optimization.
- Architecting a decentralized privacy marketplace where an LLM empowers users with granular controls, turning forced surrender into incentive-aligned, sustainable, and profitable exchanges.
- Enhancing AI chatbots, especially in healthcare AI, with uncertainty quantification, enabling statistical confidence assessment to deliver more reliable and trustworthy recommendations.

**Graduate Researcher** - *Columbia University* JUN 2020 - JUN 2021

- Optimized a Bayesian hierarchical sparse VAR model for multi-subject, multi-session fMRI analysis, leveraging high-performance computing and Gibbs sampling to enhance stability and scalability.

**Research Assistant** - *Columbia Business School* DEC 2019 - MAR 2020

- Developed a high-accuracy algorithm for identifying university-registered patents, surpassing NBER's method through data analysis, refined regular expressions, and optimized extraction in SAS.

## Teaching & Leadership

**Teaching Assistant** - *UC Los Angeles & Columbia University* JUN 2021 - PRESENT

- Instructing statistics courses, covering probability, inference, computational methods, and optimization, while mentoring students in R programming, data analysis, and reproducible workflows.

**Team Leader** - *UC San Diego* MAR 2018 - JUN 2018

- Led a team of six in a fast-paced statistical project course, driving data analysis across five real-world topics while guiding model development, decision-making, and efficient collaboration.

## Skills

- **Programming:** Python (SciPy, TensorFlow, PyTorch, rpy2), R, SQL, SAS, HPC
- **Economics:** game theory, causal inference, A/B testing, mechanism design, auctions, econometrics