Siyun (Cheryl) Yan

(804) 873-0878 | sy3199@columbia.edu | New York, NY

EDUCATION

Columbia University **New York, NY**
B.Sc. & M.Sc. in Financial Engineering Expected Dec 2026

- **GPA:** 3.8 / 4.0**; Relevant Coursework:** Machine Learning & Data Analytics, Natural Language Processing, Data Structure, Intro to Databases, Stochastic Models, Optimization Methods, Monte Carlo Simulations
- **Award**: 1st Prize in 2023 Hudson River Trading & Columbia Data Science Society Hackathon

WORK EXPERIENCE

Ask2.ai **New York, NY**
Quantitative Research Intern Jan 2026 – Present

- Expected to apply high-dimensional representation learning (e.g. R^2-tSNE, R^2-UMAP) to analyze latent structure and diagnostics of complex predictive models

Columbia Business School, Behavioral Research Lab **New York, NY**
Quantitative Research Assistant Aug 2025 – Present

- Analyzing heterogeneous treatment effects in a randomized donation campaign across 800+ individuals by applying regression-based methods and meta-learning techniques to uncover subgroup-level response patterns
- Developing parameterized RMarkdown workflows to automatically generate individualized analytic reports at scale, ensuring consistency of metrics, formatting, and interpretations across cohorts and learning teams
- Conducting cohort-level and team-level statistical analyses across Class of 2027 MBA study groups to identify behavioral patterns and distributional differences, supporting robust interpretation of survey-based results

Deloitte **New York, NY**
Data Science Intern, Columbia IEOR Partnered Industry Project Oct 2023 – Jan 2024

- Developed a panel logistic regression model (AUC = 0.84, precision = 0.78) to estimate the likelihood of gentrification onset, classifying emerging neighborhoods across Manhattan into distinct levels of growth
- Designed a relational data schema for 50K+ Yelp business records, integrating location, category, and rating data from multiple APIs through an automated ingestion pipeline (Python, SQL) for downstream ML models
- Delivered a Python-based Folium dashboard to visualize model-predicted gentrification probabilities across neighborhoods, improving high-growth area identification by 20% relative to static real-estate benchmarks

China Taiping Insurance Group, Actuarial Division **Shanghai, CN**
Quantitative Modeling Intern May 2023 – Jul 2023

- Designed a dynamic premium adjustment model that jointly captures policyholder demand (renewal probability) and claims risk through an exponential demand–risk tradeoff function to inform pricing decisions
- Simulated scenario-based policy-level P&L and aggregated results to the portfolio level, generating revenue distributions and A/E (Actual/Expected) loss metrics to evaluate pricing under alternative assumptions

HACKATHON & INDUSTRY PROGRAMS

Bloomberg Data Academy **Shanghai, CN**
Applied Data Researcher Jun 2024

- Built an ESG risk scoring model for 3,000+ APAC firms, synthesizing financial metrics and unstructured non-financial indicators across industries; identified key risk drivers using cross-sectional and longitudinal analysis
- Explored NLP-based feature extraction methods to encode qualitative ESG disclosures into model-ready representations; evaluated robustness under alternative weighting scehmes and data availability assumptions

Hudson River Trading & Columbia Data Science Society Hackathon, 1st Place Winner **New York, NY**
Data Science Group Project Oct 2023

- Developed a dynamic pricing model for the NYC subway system that adjusted fare multipliers in near real time based on hourly demand fluctuations and station-level ridership patterns across the network
- Built time-series forecasting model using ARIMA for hourly ridership prediction; implemented spectral clustering on geospatial ridership data to segment stations into unified pricing zones
- Validated revenue improvements through Monte Carlo simulations across multiple price elasticity scenarios, demonstrating consistent revenue improvements over baseline pricing in more than 95% of simulated scenarios

SKILLS

Languages: Mandarin & English; **Technical Skills:** Python, R, SQL, LaTex, Git, Matlab, Java