

Jack C. Yeung

jackyeung99@gmail.com | Portfolio | GitHub | LinkedIn

Education

Indiana University-Bloomington – M.S. in Data Science (Applied Economics) May 2026

Indiana University-Bloomington – BS in Informatics Jul 2024

Coursework: Financial Econometrics, Econometrics I & II, Machine Learning, Bayesian Statistics, Linear Algebra, Probability

Skills

Python SQL Power BI PySpark Pandas NumPy Scikit-learn PyTorch XGBoost AWS Docker Git Linux

Experience

Data Analyst Intern Jun 2025 - Aug 2025

Care Plus NJ – Paramus, NJ

- Strengthened payer negotiations by enabling accurate compliance reporting through a Power BI dashboard reconciling reported vs actual telehealth time across two data sources, improving match rates by **10%**
- Enabled reliable identification of in-network reimbursement rates by architecting a scalable PySpark pipeline processing **2B+** insurance rate records and integrating billing codes, geolocation, and taxonomy data
- Uncovered revenue opportunities through competitive market analysis of two major insurers against state averages, identifying 4 underpaid billing codes and quantifying a **26.5%** potential annual revenue uplift for senior leadership

Research Assistant (Statistical Modeling) Jun 2023 – Feb 2025

Center for Complex Networks and Systems Research (CNetS) - Bloomington, IN

- Implemented and evaluated probabilistic ranking algorithms (Bradley-Terry, Plackett-Luce) for head-to-head and multi-way ranking on real-world datasets
- Improved convergence speed on average by **14x** by developing and validating an efficient estimator extending Bradley-Terry (Newman) to the Plackett-Luce setting
- Co-authored peer-reviewed publication in *Physical Review E*, contributing to model development, implementation, and empirical validation

Publications

- [1] Jack Yeung, Daniel Kaiser, and Filippo Radicchi. Efficient inference of rankings from multibody comparisons. *Phys. Rev. E*, 112:014305.

Projects

Quantitative Trading Research Platform GitHub

Independent Study with Prof. Chang, Indiana University

- Developed systematic trading strategies using realized volatility and macroeconomic signals to dynamically allocate capital across sector ETFs
- Built modular backtesting framework to evaluate strategies using walk-forward validation and parameter sweeps (Sharpe, drawdown, turnover)
- Automated live trading by deploying a Python-based system on AWS (ECS, Docker), executing signals and rebalancing portfolios via Alpaca API

Energy Market Volatility Forecasting System Dashboard GitHub

- Built volatility forecasting pipeline using econometric models (EWMA, GARCH, HAR-RV) on **3** data sources and **50+** engineered features
- Improved forecast accuracy (QLIKE) by **25%** by engineering weather anomaly features; benchmarked models using rolling out-of-sample backtesting
- Enabled real-time monitoring by deploying daily retraining on AWS ECS/Fargate and delivering forecasts via interactive Power BI dashboards

Real Estate Pricing & Market Intelligence Platform GitHub

- Built scalable web scraping pipeline (Scrapy) to extract, validate, and normalize data for **300,000+** listings across multiple rental listing platforms
- Reduced prediction error by **55%** by engineering NLP-based features and training an XGBoost pricing model
- Identified underpriced assets and market segmentation patterns by analyzing geospatial pricing data using GeoPandas