

Jack C. Yeung

jackyeung99@gmail.com | linkedin.com/in/jackcyeung | jackyeung99.github.io

EDUCATION

Indiana University - Bloomington	May 2026
<i>Master of Science (M.S.) in Data Science, Applied Data Science – Economic Data Analytics</i>	GPA 3.95/4.0
Indiana University - Bloomington	July 2024
<i>Bachelor of Science (B.S.) in Informatics</i>	
Relevant Coursework: Financial Econometrics, Econometrics, Machine Learning, Applied Machine Learning, Bayesian Statistics, Applied Algorithms, Game Theory for Business, Engineering Cloud Computing, Exploratory Data Analysis, Performance Analytics, Network Science, Linear Algebra, Probability and Statistics	

TECHNICAL SKILLS

- Programming:** Python, SQL, HTML/CSS
Data & ML Libraries: Pandas, NumPy, Scikit-learn, PyTorch, XGBoost, NetworkX
Data Engineering & Automation: PySpark, Scrapy, Selenium, Pytest
Visualization: Matplotlib, Seaborn, Power BI
Tools & Platforms: Git, AWS, Google Cloud Platform (GCP), Docker, Linux Shell

EXPERIENCE

CarePlus New Jersey	June 2025 - August 2025
<i>Data Analyst Intern</i>	<i>Paramus, NJ</i>
○ 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	
Center for Complex Networks and Systems Research(CNetS)	June 2023 - February 2025
<i>Research Assistant</i>	<i>Bloomington, IN</i>
○ Optimized probabilistic ranking models to evaluate competitive dynamics (Formula One, elections), improving predictive performance and achieving 14x faster convergence	
○ Collaborated with a team in weekly meetings to design and refine experiments, co-authoring a research paper detailing the methodology and findings to a peer-reviewed publication in Physical Review E	

PUBLICATIONS

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

PROJECTS

Cloud-Deployed Energy Volatility Forecasting System
○ Enabled daily risk monitoring and market decisions by building an automated time-series volatility forecasting pipeline that combines financial, economic, and weather data from 3 sources and 50+ features
○ Improved forecast accuracy (QLIKE) by 25% through weather anomaly feature engineering and systematic benchmarking of econometric volatility models (EWMA, GARCH, HAR-RV) using rolling out-of-sample backtesting
○ Operationalized daily retraining and inference at scale using Docker and AWS ECS/Fargate, publishing forecasts and performance metrics to Power BI dashboards for continuous risk monitoring
Real Estate Market Intelligence & Pricing Analytics Platform
○ Designed and executed large-scale web scraping pipelines using Scrapy to extract, validate, and normalize pricing and amenity data for 300,000+ apartment units, enabling city-scale market analysis
○ Analyzed neighborhood-level pricing dynamics through geospatial EDA with GeoPandas, informing location-based investment and market segmentation strategies
○ Enabled systematic identification of underpriced rental assets by training and tuning an XGBoost pricing model incorporating NLP-derived amenity features, reducing prediction error by 55%