Chun-Yi Lee

Data Scientist

Taipei, Taiwan | chunyilee0102@gmail.com | github.com/JamesLeeCY

SUMMARY

Data scientist with a quantitative research background in statistical modelling, time-series analysis, and investment research. WorldQuant BRAIN Research Consultant (Gold level) with 65 submitted equity signals. Builds multi-agent research systems, RAG knowledge-base tools, and large-scale data pipelines that turn complex data into decision-ready reports, with stakeholder and budget experience from co-leading a US$37K nonprofit sponsorship campaign. Spent a year in the Boston area as a visiting scholar at Brandeis University, working daily in English with U.S. teams.

EXPERIENCE

Research Consultant (Conditional) | WorldQuant BRAIN Remote | Aug 2026 – Present

- Paid consulting contract developing systematic equity signals (Gold level, BRAIN Genius Program); delivered 65 signals passing quality, robustness and correlation checks (median in-sample Sharpe 1.98) across U.S., global and European markets
- Built a Python research pipeline (7,600+ simulations) screening for year-by-year stability, sub-universe robustness and correlation to existing signals; designed a multi-signal combination method under explicit correlation and crowding constraints

Visiting Scholar | Brandeis University Waltham, MA (Boston area) | Apr 2024 – Jan 2025

- Taiwan MOE-funded; coordinated the NTU–Brandeis cross-site neuroimaging data platform with U.S. collaborators

Research Data Scientist | National Taiwan University Taipei | Sep 2021 – Present

Ph.D. research, Interdisciplinary Neuroscience

- Built end-to-end Python pipelines for two fMRI studies, including a 12-week randomised controlled trial, covering ~1.9 TB of 4D imaging data (483 scan sessions) from raw DICOM through preprocessing, parcellation, and whole-brain statistical inference
- Re-engineered the permutation-testing core with an algebraic reformulation of the regression model, cutting runtime from 2.1 days to 3.3 hours (~15×) with results verified identical to machine precision
- Led a 59-participant study end to end

R&D Project Manager | UCarer (優照護) Taipei | Mar 2021 – Aug 2021

- Managed a dementia-delay program with three external partners (cognitive training, sleep, nutrition) and its cloud/API planning

Research Assistant | National Taiwan University Taipei | Mar 2019 – Aug 2021

- Built Shell/SQL data infrastructure and cross-site fMRI calibration for a Taiwan–U.S. study; first author, Biological Psychology (2021)

Substitute Military Service (Fire Service) | National Fire Agency Training Center Nantou | Apr 2018 – Mar 2019

- Hosted international delegations to the training center, mostly from Latin American and European institutions, in English

Master's Researcher | National Taiwan University Taipei | Sep 2015 – Feb 2018

- Applied statistical modelling and structured data-acquisition workflows to multi-year clinical research datasets

SELECTED PROJECTS

- **Dissertation RAG Copilot** — Citation-grounded RAG system (hybrid retrieval, ChromaDB, hallucination verification) reaching 92% Hit@1 / 100% Hit@5; applicable to internal knowledge-base search (github.com/JamesLeeCY/paper-rag-copilot)
- **LINE Group Health Triage System** — Chinese-language NLP pipeline (response delay, sentiment, escalation signals) that ranks chat groups and outputs PDF reports for managers; 94% risk-level accuracy on held-out groups of a 100-group synthetic benchmark (github.com/JamesLeeCY/line-chat-triage)
- **Precision-Weighted Multi-Agent Investment Research System** — Five-agent Taiwan-equity system tested walk-forward on a survivorship-bias-free universe (271 stocks, 8,475 predictions); a monthly revenue surprise (SUE) signal was significant in both 2019–2022 and 2023–2026 (within-sector IC +0.032 / +0.059) (github.com/JamesLeeCY/precision-weighted-investment-agent)

LEADERSHIP

- **Fundraising Co-Lead, BTBA Symposium 2025 — Boston Taiwanese Biotechnology Association (501(c)(3))** — Co-led a 6-person fundraising team for an annual two-day biotech symposium (350+ attendees); raised US$37K from 21 sponsors and donors, personally closing US$13K including the sole Diamond-tier sponsor; ran end-to-end sponsor operations (tiered packages, invoicing, payment tracking, receipts, benefit fulfilment)
- **Organizing Committee & Keynote Moderator, 2024 TAAB Annual Symposium — Taiwanese American Association of Biotechnology** — Helped organize the two-day symposium at Hunter College, New York (150–200 in-person attendees, ~30 speakers) and moderated the keynote by Dr. Ching-Yung Lin, CEO of Graphen, Inc. and former IBM Chief Scientist (Graph Computing)
- **Graduate Student President, National Taiwan University** — Represented the graduate cohort in academic and administrative coordination across departments

TECHNICAL SKILLS

Modelling: Regression and classification, time-series forecasting, hypothesis testing, factor design, walk-forward backtesting
Data & Reporting: Python (pandas, NumPy, scikit-learn), SQL, R, ETL pipelines, data-quality monitoring, dashboards and KPI reporting
AI: RAG, LLM agents, prompt engineering, PyTorch, Docker, GCP

EDUCATION

Ph.D., Interdisciplinary Neuroscience, National Taiwan University — 2021 – Expected 2027
M.S., Brain and Mind Sciences, National Taiwan University — 2015 – 2018
B.S., Forestry and Resource Conservation, National Taiwan University — 2011 – 2015
Awards: WorldQuant Summer Experience Program — Invited Participant, Certificate of Completion
Languages: Mandarin (native), English (professional working proficiency)