LAI CHUN KIT

Junior Data Analyst | Analytics Engineering (ETL, SQL, Data Modelling) & Python Internal Tools
Tanjong Malim, Perak · Willing to relocate to Klang Valley · +60 17-514 2980 · vaniepersie@gmail.com
mrbigpao.com · github.com/bigpao0516

PROFILE

Junior Data Analyst at CG Ikhlas, a Malaysian e-commerce and IT retailer, where I built its internal data platform with AI coding assistants. I build ETL pipelines for four stores: income and orders across Shopee, Lazada and TikTok Shop, plus Shopee/Lazada ads, with data models, LLM review steps and Flask tools for finance.

EDUCATION

BSc (Hons) Financial Mathematics, Universiti Malaysia Terengganu · 2022 – 2026, conferred Oct 2026 · CGPA 3.83

EXPERIENCE

Junior Data Analyst | CG Ikhlas Sdn. Bhd. (e-commerce & IT retail) | Jan 2026 – Present

- Built 12 Chrome-extension collectors with local Flask services: Shopee, Lazada and TikTok Shop income/order exports and Shopee/Lazada ads for four stores, plus competitor listings and promotions. Built staged imports with period checks, file-hash duplicate detection and seller-account checks where provided.
- Built the marketplace intelligence platform (Python/Flask, SQLite) from Jan 2026; live since Oct 2026 and used by about 10 staff, with shared sign-in for two finance apps; PostgreSQL migration in progress (outbox pattern).
- Integrated external LLM APIs for product labelling with structured-output checks: high-confidence labels auto-apply to unlabelled products; uncertain results and corrections queue for review. Commits are backed up and audited; TF-IDF + LinearSVC handles market titles without AI labels.
- Designed the data model and built a project P&L app (Flask, SQLAlchemy), live and used by staff: invoice PDFs go through text extraction, line and document total checks and confidence flags, then dated costs, four commission methods, approval and locked statements with version history.
- Built a supplier claim and credit-note app, live and used by staff: a vision model reads proofs into strict-JSON claim lines, the server checks totals and flags inconsistencies for staff review, and an AI help desk uses permission-scoped context, retries and a circuit breaker.
- AI coding assistants wrote most of the code; I owned the ideas, prompts, review and delivery. 3,700+ Python test functions; releases use Playwright, GitHub Actions and Docker health checks with rollback.
- Cut monthly reporting from 2 days to 2 hours; presented monthly ad insights to the Honeywell and ASUS Business brand teams (ASUS May–Aug 2026), supporting ad budget or campaign changes.

Data Analyst Intern | CG Ikhlas Sdn. Bhd. | Aug 2025 – Jan 2026

- Automated weekly Shopee price checks, raising throughput from about 40 SKUs a day by hand to about 250.
- Automated the Excel price comparison workbook (28 brands): XLOOKUP/LET/FILTER formulas suggest a price and action against the 3 cheapest sellers; Power Query loads promotions; 13 VBA modules merge and dedupe sheets.
- Built campaign and discount extraction with staging validation, plus dashboards, SOPs and handover notes.

Insurance Claims Specialist (part-time) | EY Auto Workshop (vehicle repair), Tanjong Malim | Jul – Oct 2024

- Prepared customer claim documents, coordinated workshop quotes and followed up on vehicle damage claims.

SKILLS

Data & languages: SQL (SQLite), Python (Pandas, SQLAlchemy, scikit-learn), JavaScript
Data engineering: ETL/ingestion pipelines, relational data modelling, staging and validation, collection work queues
Data quality & delivery: pytest, Playwright, reconciliation checks, audit trails, Git, GitHub Actions CI/CD, Docker
BI & reporting: Excel (VBA, Power Query, XLOOKUP, openpyxl), Power BI (basic), PDF/PPTX reports; ROAS, GMV
AI integration: External LLM APIs, structured-output validation, confidence gates, review queues
Apps: Flask, Jinja, Streamlit, Chrome extensions (Manifest V3), PyInstaller

PROJECTS, LANGUAGES & AVAILABILITY

Projects: Final-year research on Tesla market forecasting with news sentiment (ARIMAX vs LSTM), SRD 2025 nominee · Led the 3rd-place team in the RASIO 8.0 Data Science Competition (Universitas Padjadjaran, 2024): ASEAN economic recovery dashboard · Nomsim financial-literacy game, silver (Perak) award, Karnival Inovasi@UMT 2024
Languages: Mandarin, Cantonese, English, Malay; Japanese (about B2, self-assessed, not certified); Korean (basic)
Availability: Expected salary RM4,500–6,000 a month · start date negotiable (currently employed)
Work samples: mrbigpao.com, shared with my employer's agreement; figures redacted or invented