Luan Tran

✉ tdtruongluan1202@gmail.com · ☎ 0423538968 · Sydney NSW

LinkedIn: louistran1202
GitHub: truongluan1202
Tableau: truongluan1202

Experience

- **Data Analyst Associate** — *Sydney, Australia*
  Kpay Group — *06/2025 - Present*
  - Built scalable pipelines (**Spark, dbt, Airflow**) processing **200K+ spatial** data into **Snowflake**.
  - Segmented high-value regions and tiered hotspots using **Power BI, ArcGIS, and Mapbox**.
  - Built a tracking dashboard to monitor **funnel performance**, identify **drop-offs**, and analyse tier-level behaviour.
  - Recommended targeted strategies for each tier based on **spatial density** and **conversion patterns**.
  - Tested outreach strategies using **A/B testing**

- **Research Assistant** — *Sydney, Australia*
  Western Sydney University (Prof. Anupama Ginige) — *05/2025 - Present*
  - Built **ELT** pipelines (**dbt, Airflow, Snowflake**) for healthcare robotics data
  - Automated multi-source ingestion & transformation with optimised **SQL**, cutting latency **45%**
  - Added **dbt** tests & monitoring, reducing manual QA **60%**

- **Data Engineer Intern** — *Sydney, Australia*
  Skwill AI — *05/2024 - 03/2025*
  - Built an agentic reporting pipeline using **RAG, GPT-4o API** to generate insights from customer logs
  - Implemented **vector retrieval** and scalable **ETL** flows for fast **log indexing**
  - Automated **end-to-end** data processing, improving **reporting speed** and reducing **manual work**

- **Research Assistant** — *Sydney, Australia*
  S P Jain school of Global Management (Prof. Bilal Bahaa Zaidan) — *01/2025 - 05/2025*
  - Engineer **ML data pipeline** for fall-risk detection, **RMSE 2.35**, improved accuracy **15%**,
  - Applied **SHAP** & **k-fold validation** for explainable performance.

Education

Western Sydney University — *Sydney, Australia*
Master by Research (Computer Vision & Data Processing)
S P JAIN School of Global Management — *Sydney, Australia*
Bachelors of Data Science — **Dean's List** for **4** consecutive semesters — *Graduated, May 2025*

Certificates

- **Power BI** Data Analyst Associate (Microsoft Certified), **SQL (Advanced)** (HackerRank), **Databricks Lakehouse** Fundamentals, **AWS** Certificated Cloud Practitioner, **Data Engineer** Certificate (DataCamp), Top 100 **Shopee** Code League Finals (2022)

Technical Skills

- **Data Analytics & Engineering:** PySpark, Spark SQL, Python, SQL, PostgreSQL, NoSQL, Power BI, Tableau, Databricks, Airflow, Snowflake, MongoDB, R
- **Tools & Technologies:** AWS, GCP, Docker, CI/CD, Apache Kafka, SealSuite, Git, Scikit-learn, Langchain, Hugging Face, Databricks MLflow, PyTorch, Excel

Personal Projects

- **Calendara - AI-Powered Calendar Assistant**
  (https://ai-powered-calendar-assistant.vercel.app/)
  FastAPI, Next.js, PostgreSQL, Google Calendar API, Web Search
  - Developed full-stack conversational assistant for natural-language scheduling & event retrieval.
  - Integrated **Google Calendar API** and **web search** for contextual event management.
  - Deployed multi-provider backend (**OpenAI/Gemini**) enabling smart intent routing.

- **Airtable Clone**
  (https://airtable-clone-eta.vercel.app/)
  Next.js, TRPC, TanStack Table, PostgreSQL, Prisma
  - Built Airtable-style data grid with **virtualized** infinite scroll supporting **100k+** rows.
  - Implemented dynamic columns, filtering, and real-time collaboration via **TRPC**.
  - Optimized **PostgreSQL** queries and **Prisma ORM** for **sub-100 ms** interactions.

- **Sports Analytics Data Pipeline**
  (https://github.com/truongluan1202/Sports-Analytics-NoSQL-Database.git)
  MongoDB, PostgreSQL, Python
  - Engineered ETL pipelines processing **1.9K+** shot events from **6** European football leagues.
  - Modeled both NoSQL (**MongoDB**) and relational (**PostgreSQL**) schemas for analytical workloads.
  - Added **triggers**, validation rules, and **indexing**, improving query speed **35%** and data reliability **40%**.