Haorui Zhou

name pronounced as “how-rway” | he/they | haorui2002@gmail.com | Portfolio

EXPERIENCE

Data Acquisition Lead (Software & Hardware Infrastructure) *October 2024 – Ongoing*

Western Formula Racing, Formula SAE Team at Western University

- Architected end-to-end telemetry and metrics platform combining embedded systems, cloud infrastructure, and real-time dashboards, enabling data-driven performance optimization and diagnostics for a student racing team.
- Designed and deployed CI/CD pipelines (GitHub Actions) for automated build, test, and deployment of React dashboard to AWS (EC2, Lightsail), reducing deployment cycle time from manual process to <5 minutes.
- Built modular Python ETL services integrating CAN-bus telemetry with InfluxDB time-series database, implementing data validation, quality checks, and monitoring to ensure 24/7 system availability.
- Reduced dependency on commercial analysis tools by 80% through custom integration and API development.
- Implemented infrastructure-as-code using Docker Compose for reproducible deployments and environment consistency across development and production.

Financial Analyst Intern (Data Engineering & Automation) *May 2025 – June 2025*

CDH Investment Shanghai Office

- Automated market data collection and processing pipeline using Python Selenium scrapers and local LLM-powered summarization, reducing manual effort by 90%; designed ETL middleware for multi-source data validation, cleaning, and coordinate normalization across heterogeneous datasets.
- Deployed production PostgreSQL/PostGIS database on Dockerized Linux VM with role-based access control; authored comprehensive documentation for onboarding, schema evolution, and maintenance procedures.
- Delivered analytics reports combining statistical modeling and visualization for investment risk assessment, translating complex data into actionable business insights.

Data Engineering & Mobility Analytics Researcher August 2024 – *Ongoing*

Department of Geography and Environment at Western University

- Designed and executed quasi-experimental studies using causal inference and statistical learning to evaluate commuter behavior responses to external shocks (weather events, policy changes).
- Engineered large-scale data pipelines integrating mobility telemetry, weather APIs, and census datasets (5M+ records) for spatial-temporal modeling and simulation.
- Trained weather-conditioned WGAN models to simulate and forecast regional mobility disruptions, combining ML techniques with domain knowledge for predictive analytics.

SKILLS

- **Data Engineering & Pipelines:** Python (pandas, polars, NumPy, FastAPI), SQL (PostgreSQL, InfluxDB), ETL design, real-time streaming (WebSocket), RESTful API development, data validation & quality
- **DevOps & Infrastructure:** Docker, CI/CD (GitHub Actions), AWS (EC2, S3, Lightsail), Linux servers, network configuration, monitoring (Grafana)
- **Programming & Development:** Python (scikit-learn, pytest), R (tidyverse), C++ (basics), version control (Git)
- **Data Visualization & BI**: Grafana, React dashboards, Plotly Dash, matplotlib/seaborn
- **Statistical & Analytical Methods:** Experimental design, A/B testing, causal inference (DiD/panel models),
- **GIS & Spatial Analytics:** ArcGIS Pro, Online, Enterprise; QGIS; geopandas, PostGIS

EDUCATION

Western University, London, Ontario, Canada Expected Graduation Date: June 2026

Master of Science in Geography and Environment, Specialization in Environment and Sustainability GPA: 89.25 / 100

Focus: Data infrastructure, applied statistics, and scalable analytics pipelines for mobility and sustainability research.

UC Santa Barbara, Santa Barbara, California, the United States August 2020 – March 2024

Bachelor of Arts in Geographic Information Science, Honors Program, Distinction in the Major GPA: 3.80 / 4.0