

James Ding

604-842-6134 | james@dingjames.com | [LinkedIn](#) | [Github](#) | Portfolio

Calgary, AB • TN Visa Eligible

PROFESSIONAL SUMMARY

Data Analyst who ships production ML models, RAG pipelines, and AWS infrastructure. Looking for a team where data drives decisions

WORK EXPERIENCE

Associate Data Analyst

May 2025 – Present

AirSprint Inc.

Calgary, AB

- Conducted EDA on 12 years of flight operations data across 78K+ flights to identify key demand drivers, seasonality factors, and a COVID-era regime shift informing model design
- Built and deployed a Prophet time-series forecasting model achieving 10% weekly / 7% monthly MAPE in backtesting and 6% weekly / 2% monthly MAPE in live 2026 production, targeting 1%+ reduction in \$49M annual empty leg positioning costs
- Built automated data ingestion pipelines (FL3XX, Salesforce, Foreflight → AWS Lambda → S3 → Snowflake) with end-to-end data quality monitoring and alerting
- Designed a data-driven RBAC workaround for FL3XX's booking platform by building a cleansing and segmentation pipeline with custom business logic mapping user roles (owner, individual, full access) and a manual override layer to manage edge cases
- Built QuickSight dashboards enabling real-time monitoring of operational KPIs, flight demand trends, and fleet utilization for executive and operational stakeholders

PROJECTS

AI Judge — SCOTUS Verdict Backtester

Mar 2026

github.com/dingjamma/ai_judge

- Built a RAG backtesting pipeline over 200 landmark SCOTUS cases with temporal leakage prevention — Claude 3.5 Haiku on AWS Bedrock, FAISS vector search, MLflow experiment tracking, and a Streamlit dashboard; achieved 45.7% backtested accuracy on held-out cases

EDUCATION

Master of Data Science & Analytics

Dec 2025

University of Calgary

Calgary, AB

Bachelor of Science, Agriculture Commerce

May 2016

University of Illinois at Urbana Champaign

Champaign, IL

TECHNICAL SKILLS

Programming: Python, SQL, R, JavaScript

Machine Learning & Statistics: XGBoost, Random Forest, Prophet, ARIMA, GLM/Regression, scikit-learn, GridSearchCV, cross-validation, feature engineering, time series analysis, statistical testing, A/B testing, Experimental design, LLM application development (RAG, grounding, prompt/agent design, tool use, evaluation)

Data & Cloud: Snowflake, AWS (Lambda, S3, AppFlow, QuickSight, EventBridge, Bedrock, OpenSearch Serverless, API Gateway, CDK), pandas, NumPy, Excel, Git, REST APIs, JSON, CI/CD (Github Action)

ETL & Data Engineering: ETL/ELT pipelines, data quality monitoring, dimensional modeling, data warehousing

Visualization & BI: matplotlib, seaborn, Plotly, Plotly Dash, QuickSight, Power BI, Tableau, Streamlit, interactive dashboards