

# ROHAN NAIR

(306) 850-2044 | [rohan.nair@uwaterloo.ca](mailto:rohan.nair@uwaterloo.ca) | [linkedin.com/in/rohan-nair-108771212](https://www.linkedin.com/in/rohan-nair-108771212) | [github.com/rohannair183](https://github.com/rohannair183)

## TECHNICAL SKILLS

**Languages:** Python, SQL, R, JavaScript, Java

**Frameworks & Libraries:** Pandas, PySpark, SQLAlchemy, dbt, Scikit-learn, spaCy, Matplotlib, Seaborn, Pytest

**Technologies and Tools:** Databricks, Airflow, Spark, Kafka, BigQuery, SQL Server, PostgreSQL, Snowflake, Tableau, Looker, Metabase, Git, GitHub Actions, Azure DevOps, Docker, Terraform

## EXPERIENCE

### Data Engineer (Co-op)

January 2026 – April 2026

*Gore Mutual*

Cambridge, Ontario

- Designed and implemented a failure-handling module for Databricks pipelines that aggregated table-level errors and enforced critical-entity validation, reducing pipeline failures by **80%** across **10+** pipelines
- Designed reusable configuration-based transformation logic for Databricks pipelines, reducing reliance on complex inline SQL and standardizing data modeling patterns across datasets
- Optimized unit tests for slow-running pipeline modules, reducing CI/CD runtime by **2 hours** per day through targeted test streamlining

### Data Engineer (Co-op)

September 2024 – August 2025

*Geosyntec Consultants*

Waterloo, Ontario

- Built and maintained ETL pipelines with Python and SQL Server, transforming data from multiple sources into centralized, analytics-ready datasets across **5** projects
- Designed and implemented an automated data quality monitoring framework across **100+** SQL Server tables, enforcing completeness and validity checks that served as the data reliability foundation for downstream reporting
- Automated standardized reporting workflows integrating SQL Server and field data systems, eliminating **20+** hours of manual effort per week

### Data Analyst (Co-op)

January 2024 – April 2024

*City of Edmonton*

Edmonton, Alberta

- Designed and maintained ETL pipelines and Tableau dashboards to surface KPI metrics for business stakeholders, enabling more informed decision-making across **3** city departments
- Leveraged ESRI tools, SQL, AppSheet, and Apps Script to automate reporting workflows, reducing manual reporting overhead by **15+** hours weekly

### Software Developer (Co-op)

May 2023 – August 2023

*Audesse Automotive*

Remote

- Refactored Python codebases into standalone Cython libraries, boosting execution speed by **35%** and improving the efficiency of computationally intensive workflows
- Developed a JavaScript web application to standardize and convert vehicle log formats, enabling seamless integration with client systems and reducing compatibility issues

## PROJECTS

### Semantic News Pipeline | Python, pgvector, Gemini API, GitHub Actions

- Built a semantic news intelligence pipeline that ingests Guardian API articles, chunks and embeds text into a Supabase vector store, and uses semantic retrieval to ground Gemini-generated multi-topic briefings
- Designed YAML-driven orchestration with incremental ingestion, Parquet-based intermediate datasets, and a GitHub Actions CI/CD workflow enforcing 100% test coverage and clean linting on every change

### SEC EDGAR Insider Trading Analytics Pipeline | Python, dbt core, BigQuery, Metabase

- Built an end-to-end analytics engineering pipeline ingesting SEC EDGAR financial statement data into BigQuery using Python, modelling transformation layers in dbt (staging, intermediate, marts) with schema tests and data lineage documentation across multiple companies and fiscal quarters

## EDUCATION

### University of Waterloo

Waterloo, ON

*Bachelor of Applied Science in Management Engineering (3.85 GPA)*

2022 – 2027

- Relevant Courses: Machine Learning, Deep Learning, Database Design & Software Systems, Data Structures & Algorithms, Statistics & Probability, Stochastic Modeling, Linear Algebra, Software Engineering